

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comp-gen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:33:04 : Search time 735.169 Seconds  
(without alignments)  
10633.850 Million cell updates/sec

Title: US-09-856-681-1

Perfect score: 3093  
Sequence: 1 atgagtcagaagccttgct.....ccaatgctgcgtgacataa 3093

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Capext 1.0

Searched: 2552756 seqs, 13497:9017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

```

1: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY:HS

Result No.	Score	Query Match	Length	DB ID	Description
1	3093	100.0	3862	21	AA001233 Human semaphorin 6
2	3038	98.2	3498	21	AAA93617 Human semaphorin p
3	3032	98.0	4280	25	ABX71103 Novel human cDNA s
4	2969.6	96.0	6060	22	AAH18729 Human cDNA sequenc
5	2746.2	88.8	3550	15	AAV44295 Human secreted pro
6	2746.2	88.8	3550	22	AAF598469 Human cDNA clone c
7	2698	87.2	3333	21	AAA93618 Human semaphorin p
8	2221.8	71.8	3039	23	AA568807 DNA encoding novel

	9	2221.8	71.8	3039	23	AA568807	DNA encoding novel
10	2182.4	70.6	3041	22	AAH17625	Human cDNA sequenc	
11	1881.6	60.8	1890	21	AA93630	Human semaphorin p	
12	1824.4	59.0	2293	22	AAH15834	Human cDNA sequenc	
13	1742.4	56.3	2227	22	AAH94365	Human full-length	
14	1666.4	53.9	2306	22	AAH17861	Human cDNA sequenc	
15	1592	51.5	2592	23	AA568253	DNA encoding novel	
16	1462	47.3	2123	22	AAH13955	Human cDNA sequenc	
17	1216	39.3	2262	24	ABK34739	Human cDNA for nov	
18	962	31.1	1054	20	AAH84066	Human ZSMF-3 codin	
19	808.4	26.1	1047	20	AAH84076	Mouse ZSMF-3 codin	
20	788	25.5	839	22	AAH05402	Human cDNA clone (	
21	758.6	24.5	814	22	AAK92327	Human cDNA 5'-end	
22	758.6	24.5	814	22	AAK93383	Human cDNA clone r	
23	751.2	24.3	780	22	AAH05580	Human cDNA clone (	
24	742.6	24.0	788	22	AAK93916	Human cDNA clone r	
25	708.8	22.9	963	20	AAH84067	Degenerate Human Z	
26	686.8	22.2	963	20	AAH84077	Degenerate Mouse Z	
27	669	21.6	4234	24	ABA00055	CADHP-2 coding seq	
28	651.6	21.4	3205	24	AA564380	Human semaphorin 6	
29	657.4	21.3	662	22	AAH08370	Human cDNA clone (	
30	657.4	21.3	1923	22	AAH42598	Partial cDNA sequenc	
31	657.4	21.3	3694	22	AAH42597	Nucleotide sequenc	
32	642	20.8	2191	24	AB564381	Human semaphorin-i	
33	642	20.8	2359	24	AB564383	Human semaphorin-1	
34	642	20.8	3196	24	AB564382	Human semaphorin-1	
35	642	20.8	3364	24	AB564384	Human semaphorin-1	
36	640.6	20.7	889	23	AA568806	DNA encoding novel	
37	577.4	18.7	1896	24	AB564379	Human semaphorin-1	
38	577.4	18.7	2014	24	AAH38696	Human LP221 secret	
39	570.8	18.5	786	22	AAH05233	Human cDNA clone (	
40	566.4	18.3	6645	25	ABX34714	Human mdmt cDNA SE	
41	565	18.3	1472	21	AAH98050	Human colon cancer	
42	536.6	18.0	2506	25	AB235865	Human secretory po	
43	535.4	18.0	3524	18	AAH72108	Human semaphorin 2	
44	535.4	18.0	3524	25	AAH08822	Angiogenesis-assoc	
45	533.8	17.9	3721	21	AAA96343	cDNA encoding a no	

# ALIGNMENTS

## RESULT 1

AA001233  
ID AAD01233 standard; DNA: 3862 BP.

AC AAD01233;

DT 04-OCT-2000 (first entry)

DE Human semaphorin 6A-1 cDNA.

KW Human semaphorin 6A-1; (HSA)SEMA6A-1; neuronal development; apoptosis;  
neuronal regeneration; Fna/VASP protein family; immunomodulatory;  
KW gene therapy; diagnostic agent; therapeutic agent; differentiation;  
KW cytoskeletal stabilisation; plasticity; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 658...3750

FT /\*tag= a

FT /\*product= "Human semaphorin 6A-1"

FT /\*note= "This region is specifically claimed as

FT SEQ ID NO: 1 in claim 1"

FT misc\_feature 3532..3747

FT /\*tag= b

FT /\*note= "Encodes the binding domain of semaphorin

FT 6A-1 which selectively binds to members of Fna/VASP

FT protein family"

PN WO200031252-A1.

XX

PD 02-JUN-2000.  
XX  
PF 26-NOV-1999; 99W0-EP09215.  
XX  
PR 26-NOV-1998; 98EP-0122441.  
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PA  
XX Behl C, Klostermann A;  
XX WPI: 2000-400065/34.  
XX P-PSDB; AAY71460.  
XX  
XX Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,  
XX therapeutic agent, for modulating immune system, in gene therapy or for  
XX effecting differentiation, cytoskeletal stabilization and/or plasticity  
XX  
XX Claim 1; Fig 2; 53pp; English.  
XX  
XX The present sequence is a cDNA encoding transmembranous human semaphorin  
XX 6A-1 (HSA)SEMA6A-1) which is involved in neuronal development and  
XX regeneration mechanisms during apoptosis. Semaphorin is a family of  
XX proteins displaying secreted or transmembrane-based repulsive guidance  
XX cues critically involved in neuronal development. The present sequence  
XX was isolated from human 1-ZAP Express cDNA library which was screened  
XX using a PCR fragment amplified from human neuroblastoma cell line  
XX SK-N-MC. cDNA. The (HSA)SEMA6A-1 protein contains a Zyxin-like domain  
XX that selectively binds to members of Ena/VASP protein family especially  
XX Evi. Expression of (HSA)SEMA6A-1 is highest in embryonic brain and  
XX kidney and moderate in lung. The present sequence is useful as diagnostic  
XX and therapeutic agents, for modulating the immune system, in gene  
XX therapy, for effecting differentiation, cytoskeletal stabilisation  
XX and plasticity.  
XX  
XX Sequence 3962 BP; 971 A; 1211 C; 967 G; 813 T; 0 other;  
SQ  
Query Match 100.0%; Score 3093; DB 21; Length 3962;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3093; Conservative 0; Mismatches 0; Gaps 0;  
QY 1 ATGAGTCAGAGACCTTCGTCGCTATATTTACACATGCTACACCTTGTGGGCTGCTTC 60  
DB 658 ATGAGTCAGAGACCTTCGTCGCTATATTTACACATGCTACACCTTGTGGGCTGCTTC 717  
QY 61 CCAGAGATTCAGCCCAATCAGTATTTCCGATCGCAACTATACAAACACAGTATCCCGTG 129  
DB 718 CCAGAGATTCAGCCCAATCAGTATTTCCGATCGCAACTATACAAACACAGTATCCCGTG 777  
QY 121 TTGTGGGCCACAGCCAGCGGAGAACACACACAGAGCCAGCGCTGGACATCCAGATG 180  
DB 778 TTGTGGGCCACAGCCAGCGGAGAACACACAGAGCCAGCGCTGGACATCCAGATG 837  
QY 181 ATTATGATCATGACAGGAACCTCTACATTCCTCTAGGGACCATATTTATCTGTTGAT 240  
DB 838 ATTATGATCATGACAGGAACCTCTACATTCCTCTAGGGACCATATTTATCTGTTGAT 897  
QY 241 ATAGACATCATCACAGCGAGAAATTTATTTAGCAAAAAACATGACATGAAATCAGA 300  
DB 898 ATAGACATCATCACAGCGAGAAATTTATTTAGCAAAAAACATGACATGAAATCAGA 957  
QY 301 CAGGCGGATGACACATGACAGATGAAGGAAACATAGAGTATGAGTCCCAACCTT 360  
DB 958 CAGGCGGATGACACATGACAGATGAAGGAAACATAGAGTATGAGTCCCAACCTT 1017  
QY 361 ATTAAGTTCCTTAAGAAAAACGATGATGCATTTGTTTGTGGAACCTTAATGCGCTTC 420  
DB 1016 ATTAAGTTCCTTAAGAAAAACGATGATGCATTTGTTTGTGGAACCTTAATGCGCTTC 1077  
QY 421 ACCCTTCTCGAGAACTATAGATGATATGATGATGATGATGATGATGATGATGATGAT 480  
DB 1078 ACCCTTCTCGAGAACTATAGATGATATGATGATGATGATGATGATGATGATGATGAT 1137

QY 481 GGAATGGCCAGATGCCCATATGATGCCAAACATGCCCAACCTTGCACCTGTTTGCAGATGGA 540  
DB 1138 GGAATGGCCAGATGCCCATATGATGCCAAACATGCCCAACCTTGCACCTGTTTGCAGATGGA 1197  
QY 541 AAACATATATCTAGCCACAGTCTGCTTCCCTTCCCATTTGACGAGTCAATTTACCGAGT 600  
DB 1198 AAACATATATCTAGCCACAGTCTGCTTCCCTTCCCATTTGACGAGTCAATTTACCGAGT 1257  
QY 601 CTTCGAGAAAGCCCTACCCCTGGGACCGCTCAAGCAGCATTCACAAATGTTTGAAGAACCA 660  
DB 1258 CTTCGAGAAAGCCCTACCCCTGGGACCGCTCAAGCAGCATTCACAAATGTTTGAAGAACCA 1317  
QY 661 TACTTGTTCAGCGCTGGATTACGGAGATATATCTACTTCTTCTTCACGGAAATAGCA 720  
DB 1318 TACTTGTTCAGCGCTGGATTACGGAGATATATCTACTTCTTCTTCACGGAAATAGCA 1377  
QY 721 GTGGAGTATACACCATCGGAAAGTGTGTTTCCCAAGAGTGGCTCAGGTTTGTAAAGAT 780  
DB 1378 GTGGAGTATACACCATCGGAAAGTGTGTTTCCCAAGAGTGGCTCAGGTTTGTAAAGAT 1437  
QY 781 GATATGGGAGTATCAAAAGAGTCTCGAGAAACAGTGGAGTCTCTCCTGAAGCGCGC 840  
DB 1438 GATATGGGAGTATCAAAAGAGTCTCGAGAAACAGTGGAGTCTCTCCTGAAGCGCGC 1497  
QY 841 TTGAACCTCTCAGTTCCTGGAGACTCTCAITTTTATTTCAACATTCCTCCAGGCAATTACA 900  
DB 1498 TTGAACCTCTCAGTTCCTGGAGACTCTCAITTTTATTTCAACATTCCTCCAGGCAATTACA 1557  
QY 901 GATGTGATTCGTATCAAGCGGCGTGATGTTCTCTGGCAAGTCTTCTACACCTTATAAC 960  
DB 1558 GATGTGATTCGTATCAAGCGGCGTGATGTTCTCTGGCAAGTCTTCTACACCTTATAAC 1617  
QY 961 ACCATCCCTGGCTGCGAGTCTGTCCTAIGACATGCTTGACATTCGCGAGTGTCTTACT 1020  
DB 1618 ACCATCCCTGGCTGCGAGTCTGTCCTAIGACATGCTTGACATTCGCGAGTGTCTTACT 1677  
QY 1021 GGGAGATTCAGGAACAGAGTCTCTGATTCACCTGGACACAGTCTCCTGATGAACGA 1080  
DB 1678 GGGAGATTCAGGAACAGAGTCTCTGATTCACCTGGACACAGTCTCCTGATGAACGA 1737  
QY 1081 GTTCTCAAGCCAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 1738 GTTCTCAAGCCAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797  
QY 1141 TCCATGATGTTCCCTGATGATACCCCTGATGATGATGATGATGATGATGATGATGATGATG 1200  
DB 1798 TCCATGATGTTCCCTGATGATACCCCTGATGATGATGATGATGATGATGATGATGATGATG 1857  
QY 1201 GCAGTGGCTCTCTCTCAAGCGGCGTGTGCTGAGAACATGCTGAGTACCGGCTT 1260  
DB 1858 GCAGTGGCTCTCTCTCAAGCGGCGTGTGCTGAGAACATGCTGAGTACCGGCTT 1917  
QY 1261 ACCAAATTCAGTGGACACAGTCTGCTGGGCGCATATCAGAAATGAGTGTGTTTCTG 1320  
DB 1918 ACCAAATTCAGTGGACACAGTCTGCTGGGCGCATATCAGAAATGAGTGTGTTTCTG 1977  
QY 1321 GGATCAGAGAGGAATCATCTTGAAGTTTGTGGCCAGATAGGAATAGTGTGTTTCTA 1380  
DB 1978 GGATCAGAGAGGAATCATCTTGAAGTTTGTGGCCAGATAGGAATAGTGTGTTTCTA 2037  
QY 1381 AATGACAGCTTTTCTCTGGAGGAGATGATGTTTACAACTCTGAAATGACGCTATGAT 1440  
DB 2038 AATGACAGCTTTTCTCTGGAGGAGATGATGTTTACAACTCTGAAATGACGCTATGAT 2097  
QY 1441 GGAGTGGAGACAAAGGATCATGGGCATGCGAGCTGGACAGCAAGCAGCTCTCTGTAT 1500  
DB 2098 GGAGTGGAGACAAAGGATCATGGGCATGCGAGCTGGACAGCAAGCAGCTCTCTGTAT 2157  
QY 1501 GTTGGGTCTCTACCTGTGTATAGGTTCCCTTGGCCCGTGTGACGACATGGGAAG 1560  
DB 2158 GTTGGGTCTCTACCTGTGTATAGGTTCCCTTGGCCCGTGTGACGACATGGGAAG 2217  
QY 1561 TGTAAAAAACCTGTATTGCTCTCCAGAGACCCATATTCTGATGATGAAGAGGAGTGGT 1620



Db 2218 TCTAAAGAACCTGATTGCTCCAGAGCCATATTGGATGATCAAGAGAGGTGGT 2277  
QY 1621 GCTTCAGCATTATATCCACCAAGCAGACACTGCTTTTGAGCAGGACATAGAGCGTGGC 1680  
Db 2278 GCTTCAGCATTATATCCACCAAGCAGACACTGCTTTTGAGCAGGACATAGAGCGTGGC 2337  
QY 1681 ATACAGATGGCTTGGGAGCTGTCAAAATGCTTTTGAGCAGTGAATGGGCATTCAGT 1740  
Db 2338 AATACAGATGGCTTGGGAGCTGTCAAAATGCTTTTGAGCAGTGAATGGGCATTCAGT 2397  
QY 1741 TCCCTTCGCCAGCAGACACATACAGATTCAGCGCTCAAGAGGGGTATGAGTCTAGG 1806  
Db 2398 TCCCTTCGCCAGCAGACACATACAGATTCAGCGCTCAAGAGGGGTATGAGTCTAGG 2457  
QY 1801 GGAGGAATGCTGAGCTGAGAGCATCTGCTGACTCACTGACCTGACAGACAGACCTTTGGG 1860  
Db 2458 GGAGGAATGCTGAGCTGAGAGCATCTGCTGACTCACTGACCTGACAGACAGACCTTTGGG 2517  
QY 1861 GCAGTCTCTCCCAATAATCACTAAGACAGAGAGGAGTGAATCGGGAAGTACTGANA 1920  
Db 2518 GCAGTCTCTCCCAATAATCACTAAGACAGAGAGGAGTGAATCGGGAAGTACTGANA 2577  
QY 1921 GGCACGACAGCTGCTCCCTGCTACCCCTCTTGGGCAATGCACTGCTGCTTCTGCTC 1980  
Db 2578 GGCACGACAGCTGCTCCCTGCTACCCCTCTTGGGCAATGCACTGCTGCTTCTGCTC 2537  
QY 1981 ATGGGCGCGTCTTCGGGATACAGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
Db 2638 ATGGGCGCGTCTTCCTGCGGATACAGCTGCTACTGCTGCTGCTGCTGCTGCTGCTG 2697  
QY 2041 GTGGCTGTGGTCAAGGAGAGAGGAGCTCACCCACTCGCGCGGGGTCCCATGAGC 2100  
Db 2698 GTGGCTGTGGTCAAGGAGAGAGGAGCTCACCCACTCGCGCGGGGTCCCATGAGC 2757  
QY 2101 AGGCTCACCAGCTCAGCGGCTCTTTTGGGACATCAATCCAAAGACCCAAAGCGGAG 2160  
Db 2758 AGGCTCACCAGCTCAGCGGCTCTTTTGGGACATCAATCCAAAGACCCAAAGCGGAG 2817  
QY 2151 GCATCCTCAGCCACTCATGACACAGCAGAGCTCGGCACACTCCCGGCAACAGCGCCAG 2220  
Db 2818 GCATCCTCAGCCACTCATGACACAGCAGAGCTCGGCACACTCCCGGCAACAGCGCCAG 2877  
QY 2221 ATGCTCATTAAGCAGACACACACCTGGACCTGAGCGGCTCCCGACCCAGAGTCA 2280  
Db 2878 ATGCTCATTAAGCAGACACACACCTGGACCTGAGCGGCTCCCGACCCAGAGTCA 2937  
QY 2281 ACCCAAGCTGACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
Db 2938 ACCCAAGCTGACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2997  
QY 2341 CAGAACCTCATCAATGCTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
Db 2998 CAGAACCTCATCAATGCTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3057  
QY 2401 ACGGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 2460  
Db 3058 ACGGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 3117  
QY 2461 ACCGACAGGGCTACGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520  
Db 3118 ACCGACAGGGCTACGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3177  
QY 2521 ATGGCGCTGGAGGACAGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2580  
Db 3178 ATGGCGCTGGAGGACAGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 3237  
QY 2581 AGCAAGAGTCCCAAGCATGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2640  
Db 3238 AGCAAGAGTCCCAAGCATGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 3297  
QY 2641 GTTCCACAGCGGAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700

Db 3296 GTTCCACAGCGGAGGAGGCTTCCCTGGTCCCGGGAGGCTTCCCTGTCTCAGACCGGTCTA 3357  
QY 2701 AGCAAGCGCTGAAATGCAACACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2760  
Db 3358 AGCAAGCGCTGAAATGCAACACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3417  
QY 2761 CCACAGAACTCTGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Db 3418 CCACAGAACTCTGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3477  
QY 2821 TCCCTCAATTCCTCTCACTCTCCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Db 3478 TCCCTCAATTCCTCTCACTCTCCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3537  
QY 2881 CCGGCGCGGAG 2940  
Db 3538 CCGGCGCGGAG 3597  
QY 2941 GTGACTGTCTGAGGAG 3000  
Db 3598 GTGACTGTCTGAGGAG 3657  
QY 3001 AAGCTGAGGCTGCTGCTTAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Db 3658 AAGCTGAGGCTGCTGCTTAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3717  
QY 3061 ACATCAATGAAGCCCAATGAAGGCTGATCAATA 3093  
Db 3718 ACATCAATGAAGCCCAATGAAGGCTGATCAATA 3750

RESULT 2  
AAA93617  
ID AAA93617 standard; DNA; 3498 BP.  
XX  
AC AAA93617;  
XX  
XT 16-JAN-2001 (first entry)  
XX  
DE Human semaphorin protein-like splice variant SECX 2864933-1 DNA.  
XX  
KW SECX protein; human; secreted; membrane-associated; cancer;  
KW proliferation regulator; differentiation regulator; non-malignant tumour;  
KW immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
KW infection; inflammatory disorder; arthritis; hematopoietic disorder;  
KW skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
KW neurological disease; Alzheimer's disease; trauma; wounding;  
KW spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
KW anti-HIV; anti-inflammatory; antiarthritic; antiarteriosclerotic;  
KW neuroprotective; vulnery; antiallergic; antimicrobial; cardiant;  
KW dermatological; gene therapy; ds.  
XX  
CS Homo sapiens.  
XX  
PN W0200053742-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-US06280.  
XX  
PR 09-MAR-1999; 99US-0123667.  
XX  
PR 08-MAR-2000; 2000US-0123667.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinketsu RA;  
XX  
DR WPI: 2000-5943318/56.  
XX  
XX P-PSDB: AAB23030.  
XX  
PT Novel human membrane associated or secreted polypeptides and  
PT polynucleotides useful for diagnosis, prevention and treatment of  
PT pathological states such as cancer, immune, cardiovascular and

PT	neurological disorders -	QY	541	AACTATATCTACGCCACAGTGA	CTCTTGGCATGACCGAGTCA	TTTACCAGCT	600
XX	Claim 3; Fig 2; 151pp; English.	DB	754	AACTATATCTACGCCACAGTGA	CTCTTGGCATGACCGAGTCA	TTTACCAGCT	813
XX		QY	601	CTTGGAGAAAGCCCTACGCTCG	CGGCGGTCAGACAGATTC	AAATGGTGAAGAACCA	660
CC	Sequences AA93616-A93631 and AA93673-A93676 represent nucleic acids	DB	814	CTTGGAGAAAGCCCTACGCTCG	CGGCGGTCAGACAGATTC	AAATGGTGAAGAACCA	873
CC	which encode human SECX proteins (AA93629-A93648). The SECX proteins	QY	661	TACTTTGTTCAAAGCCGTGGGAT	TACGGAGATTAATCTAC	TTCTTCTTCAGGGAAATAGCA	720
CC	of the invention are either secreted or membrane-associated proteins	DB	874	TACTTTGTTCAAAGCCGTGGGAT	TACGGAGATTAATCTAC	TTCTTCTTCAGGGAAATAGCA	933
CC	and act as regulator of cellular proliferation and differentiation. SECX	QY	721	GTGAGATATACACATCGGAAAG	GTAGTTTCCCAAGAGTGGCT	CAGGTTTGTGAGAAAT	780
CC	proteins or nucleotides are useful for diagnosing the presence of, or	DB	934	GTGAGATATACACATCGGAAAG	GTAGTTTCCCAAGAGTGGCT	CAGGTTTGTGAGAAAT	993
CC	predisposition to, a disease associated with altered levels of SECX	QY	781	GATATGGAGAGATCTCAAGAG	TCCTCGAGAAACAGTGGAC	TCGTTCCCTGAAGGCGCGC	840
CC	proteins and nucleotides. The SECX proteins are also useful to screen	DB	994	GATATGGAGAGATCTCAAGAG	TCCTCGAGAAACAGTGGAC	TCGTTCCCTGAAGGCGCGC	1053
CC	compounds that modulate SECX activity or expression. The interaction of	QY	841	TTGAATGCTCAGTTCCTGAGAG	CTCTCATTTTTATTTCAACAT	TCTCCAGGCGAGTTACA	900
CC	a SECX protein with other cellular proteins may be useful to modulate	DB	1054	TTGAATGCTCAGTTCCTGAGAG	CTCTCATTTTTATTTCAACAT	TCTCCAGGCGAGTTACA	1113
CC	the activity of a partner protein. Cellular proliferation, cellular	QY	901	GATGATGCTATCAACGGGCGT	GTATGTTCTCTGGCAACGTT	TTTCTACACCTTATAAC	960
CC	differentiation and cell survival. SECX nucleotides are useful for the	DB	1114	GATGATGCTATCAACGGGCGT	GTATGTTCTCTGGCAACGTT	TTTCTACACCTTATAAC	1173
CC	recombinant expression of SECX protein, and may be used to modulate	QY	961	AGCATCCCTGGGTCCTGAGTCT	GTGGCTATGACATGCTGACAT	TGCGGAGTGTTTTACT	1020
CC	or genetic lesions in the SECX gene. They may also be used to modulate	DB	1174	AGCATCCCTGGGTCCTGAGTCT	GTGGCTATGACATGCTGACAT	TGCGGAGTGTTTTACT	1233
CC	SECX expression (e.g., using antisense oligonucleotides). SECX nucleic	QY	1021	GGGAGATTCAGAGAACAGAGT	CTCTTGATTCACATGGACAC	AGTTCCTGATGAACGA	1080
CC	acid sequences are also useful for identifying a cell or tissue type in	DB	1234	GGGAGATTCAGAGAACAGAGT	CTCTTGATTCACATGGACAC	AGTTCCTGATGAACGA	1293
CC	a biological sample, and in forensic biology. SECX primers or probes are	QY	1081	GTTCCTAAAGCCAGGCGAGGTT	GTCTGGCTGGCTTCTCTAG	AAAGATATGAACAC	1140
CC	useful for detecting the presence of SECX nucleotides and for screening	DB	1294	GTTCCTAAAGCCAGGCGAGGTT	GTCTGGCTGGCTTCTCTAG	AAAGATATGAACAC	1353
CC	tissue cultures for contamination. Diseases that may be treated or	QY	1141	TCCATGATGTTCCCTGTGAT	GCATACCTGAACTCATCAAG	CGGACCGGCTCATGGATGAG	1200
CC	prevented using SECX proteins or nucleotides include cancer (e.g.,	DB	1354	TCCATGATGTTCCCTGTGAT	GCATACCTGAACTCATCAAG	CGGACCGGCTCATGGATGAG	1413
CC	colorectal carcinoma, prostate cancer), benign tumours, immune disorders	QY	1201	GCAGTGCCTCTCATCTTCAAC	AGGCCATGGTTCTGAGACAA	TATGGTTCAGATACCGGCTT	1260
CC	(including autoimmune diseases, transplant rejection, allergies, AIDS),	DB	1414	GCAGTGCCTCTCATCTTCAAC	AGGCCATGGTTCTGAGACAA	TATGGTTCAGATACCGGCTT	1473
CC	infections, inflammatory disorders, arthritis, haematopoietic disorders,	QY	1261	ACCAAAATTCAGTGGACACAC	AGCTGCTGGGCCATATCAGA	ATCACACTGTGTTTCTG	1320
CC	skin disorders, cardiovascular disorders, atherosclerosis, restenosis,	DB	1474	ACCAAAATTCAGTGGACACAC	AGCTGCTGGGCCATATCAGA	ATCACACTGTGTTTCTG	1593
CC	neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,	QY	1321	GGATCAGAGAGGGAATCATCTT	GAAAGTTTGTGGCCAGAA	TAGAAATAGTGGTTTCTA	1380
CC	surgical or traumatic wounds, spinal cord injury), and skeletal	DB	1593	GGATCAGAGAGGGAATCATCTT	GAAAGTTTGTGGCCAGAA	TAGAAATAGTGGTTTCTA	1593
CC	disorders.	QY	1381	RATGACGCTTTTCCCTGGAG	GAGATGAGTGTTCACACT	CTGAAAATCAGCTATGAT	1440
XX	Sequence 3498 BP; 917 A; 966 C; 859 G; 725 T; 1 other:	DB	1594	RATGACGCTTTTCCCTGGAG	GAGATGAGTGTTCACACT	CTGAAAATCAGCTATGAT	1553
QY	Query Match 98.2%; Score 3036; DB 21; Length 3498;	QY	1441	GGAGTCGAAGACAAAGSATCAT	TGGGCATGCGAGCAGAG	CAAGCAAGCTCTCTGTAT	1500
DB	Best Local Similarity 99.3%; Freq. No. 0;	DB	1654	GGAGTCGAAGACAAAGSATCAT	TGGGCATGCGAGCAGAG	CAAGCAAGCTCTCTGTAT	1713
Matches 3072; Conservative 0; Mismatches 21; Indels 2; Gaps 2;		QY	1501	GTTCGGTCTCTACCTGTGTAT	AAAGGTTCCCTTGGCCGGT	GTGACACATGGAAG	1560
QY	1 ATGAGGTCAGAGGCTTGTGCTATATTTTACACTGCTACACTTGTGCTGGGCTGTTTC	DB	1714	GTTCGGTCTCTACCTGTGTAT	AAAGGTTCCCTTGGCCGGT	GTGACACATGGAAG	1773
DB	214 ATGAGGTCAGAGGCTTGTGCTATATTTACACTGCTACACTTGTGCTGGGCTGTTTC	QY	1561	TGTAAGGCTCTATTCCTCCAG	AGACCCCATATTGTGGAT	GGATAAAGGAGTGGT	1620
QY	61 CCAAGATTTCTGAGCAATCAGTATTTCCGATGGCAACTATACAAACAGATTCGGTG	DB	1774	TGTAAGGCTCTATTCCTCCAG	AGACCCCATATTGTGGAT	GGATAAAGGAGTGGT	1833
DB	274 CCAGAGATTTCTGAGCCATCAGTATTTCCGATGGCAACTATACAAACAGATTCGGTG	QY	1621	GCCTGACGCCATTTATACCA	CCCAACAGCAGACTGACT	TTTGAGCAGGACATAGAGCGTGGC	1680
QY	121 TTTTGGGCCACAAAGCCAGGAGGAAACCCACAGAGGACAGGCTGGACATCCAGATG						
DB	334 TTTTGGGCCACAAAGCCAGGAGGAAACCCACAGAGGACAGGCTGGACATCCAGATG						
QY	181 ATTATGATCATGACGGAACCTCTACATGCTGCTAGGACCATATTTATCTGTTGAT						
DB	394 ATTATGATCATGACGGAACCTCTACATGCTGCTAGGACCATATTTATCTGTTGAT						
QY	241 ATAGACATCATCACACGGAAATATTTATGTAGCAAAATCTGACATGGAATCTAGA						
DB	454 ATAGACATCATCACACGGAAATATTTATGTAGCAAAATCTGACATGGAATCTAGA						
QY	301 CAGGCGGATGTAGACATGCAGATGAAGGAAACATAGGATGAGTGCACACTT						
DB	514 CAGGCGGATGTAGACATGCAGATGAAGGAAACATAGGATGAGTGCACACTT						
QY	361 ATTAAGTCTCTTAAAGAAAAGGATGATGCTATTTGTGTGGAACATATGCTTC						
DB	574 ATTAAGTCTCTTAAAGAAAAGGATGATGCTATTTGTGTGGAACATATGCTTC						
QY	421 AACCTTCTCGAGAACTATAGATGATACATTTGGAACCATTCGGGGATGATTCAGC						
DB	634 AACCTTCTCGAGAACTATAGATGATGATACATTTGGAACCATTCGGGGATGATTCAGC						
QY	481 GGAATGGCCAGATGCCCATATATGTCACCAACATGCCAAGCTTGCACCTGTTTGAGATGGA						
DB	694 GGAATGGCCAGATGCCCATATATGTCACCAACATGCCAAGCTTGCACCTGTTTGAGATGGA						

1834	DB	GCCTCAGCCATTTATCACCACAGCAGACTGACTTTTGACAGGACATAGAGCGTGGC	1893
1881	QY	AATACAGATGGCTGGGGACATGTCACAAATTCCTTTGTGGCACTGAATGGGCATTCACAT	1740
1894	DB	AATACAGATGGCTGGGGACATGTCACAAATTCCTTTGTGGCACTGAATGGGCATTCACAT	1953
1741	QY	TCGCTCTTCCAGACAAACACATCAGATTCGAGGGCTCAAGAGGGTATGAGTCTAGG	1800
1954	DB	TCGCTCTTCCAGACAAACACATCAGATTCGAGGGCTCAAGAGGGTATGAGTCTAGG	2013
1801	QY	GGAGAACTGCTGGAGTGGAGCATCTGCTTGACTCACCCTGACAGACAGACGCTTGGGG	1860
2014	DB	GGAGAACTGCTGGAGTGGAGCATCTGCTTGACTCACCCTGACAGACAGACGCTTGGGG	2073
1861	QY	GGAGTGTCTCCCATAAATCACCAGACAAAGAGGAGTGAATCGGGAAGATTACCTCAAA	1920
2074	DB	GGAGTGTCTCCCATAAATCACCAGACAAAGAGGAGTGAATCGGGAAGATTACCTCAAA	2133
1921	QY	GGCCAGCACAGCTGGTTCGGCTCACGCTCTTGGGCATTCGAGTCAATCCTGGCTTCGCTC	1980
2134	DB	GGCCAGCACAGCTGGTTCGGCTCACGCTCTTGGGCATTCGAGTCAATCCTGGCTTCGCTC	2193
1981	QY	ATGGGGCGGTCTCTTCGGGCATCACCGTCTACTGGTCTGTGCAATCATCGGGGCCAAGAC	2040
2194	DB	ATGGGGCGGTCTCTTCGGGCATCACCGTCTACTGGTCTGTGCAATCATCGGGGCCAAGAC	2253
2041	QY	GTGGCTGTGTCAGCGCAAGAGAGAGAGCTCACCCACTCGCGGGGGTCCATGAGC	2100
2254	DB	GTGGCTGTGTCAGCGCAAGAGAGAGAGCTCACCCACTCGCGGGGGTCCATGAGC	2313
2101	QY	ACGGTCAACCAAGCTCAGCGGCTCTTTTGGGGACACTCAATTCAAAGACCCAAAGCGCAG	2150
2314	DB	ACGGTCAACCAAGCTCAGCGGCTCTTTTGGGGACACTCAATTCAAAGACCCAAAGCGCAG	2373
2161	QY	GGCATCCTCAGCGCATCATGCAACAGCGGAGCTGCGCACTCCCGGCAACAGCGCCAG	2220
2374	DB	GGCATCCTCAGCGCATCATGCAACAGCGGAGCTGCGCACTCCCGGCAACAGCGCCAG	2433
2221	QY	ATGCTCATTAAGACACACAGCACCACCTGGAGCTCAGCGGCTCCCGCCACCCAGAGTCA	2280
2434	DB	ATGCTCATTAAGACACACAGCACCACCTGGAGCTCAGCGGCTCCCGCCACCCAGAGTCA	2493
2281	QY	ACCCCAACGCTGCAGCAGAAACGGGAACCCAGCGCGGAGCGCGAGTGGGAGAGGAC	2340
2494	DB	ACCCCAACGCTGCAGCAGAAACGGGAACCCAGCGCGGAGCGCGAGTGGGAGAGGAC	2553
2341	QY	CAGACCTCATCAATGCTGTGCAAAAGGACATCGCGGCCCATGGCTCGCTGTGATTGCG	2400
2554	DB	CAGACCTCATCAATGCTGTGCAAAAGGACATCGCGGCCCATGGCTCGCTGTGATTGCG	2613
2401	QY	ACGGAGCTGCCCTGGGGGCTTCCCGCAGCCATATCCCAAGCTGTGTGTCTGGCCATC	2460
2614	DB	ACGGAGCTGCCCTGGGGGCTTCCCGCAGCCATATCCCAAGCTGTGTGTCTGGCCATC	2673
2461	QY	ACGCACAGGGCTACCAAGATAGTAGTGGACGACGCGCAAAATGACGAGGTGGCGCAG	2520
2674	DB	ACGCACAGGGCTACCAAGATAGTAGTGGACGACGCGCAAAATGACGAGGTGGCGCAG	2733
2521	QY	ATGGCGCTGGAGGACCAAGCGCGGCACACTGAGATTAAGACCATCAAGGAACAATCAGC	2580
2734	DB	ATGGCGCTGGAGGACCAAGCGCGGCACACTGAGATTAAGACCATCAAGGAACAATCAGC	2793
2581	QY	AGCAAGTCTCCCAACCATGGGTGAACCTTGTGCAAACTGACAGGCTGCGCCGCCAAA	2640
2794	DB	AGCAAGTCTCCCAACCATGGGTGAACCTTGTGCAAACTGACAGGCTGCGCCGCCAAA	2853
2641	QY	GTTCACAGCGGGAGGCTCCCTGGGTCCCGCGGGAGCTCCCTGTCTCAGACCGGCTA	2700
2854	DB	GTTCACAGCGGGAGGCTCCCTGGGTCCCGCGGGAGCTCCCTGTCTCAGACCGGCTA	2913
2701	QY	AGCAAGCGGTGCAAAATGCATCACTCTCTCTCTACCGGGTTCACATAAGAGAGCTAC	2760

2914	ACGACAGCGGGTGGAAATGCACCACCTCCTCTCTTCACGGGGTTGACTATTAAGAGGAGCTAC	2878
QY	CCACAGCAACTCGCTCAGGAGAAGGCCACACGGCCACC-ACCTCTCAAAGAAAACAACACTAA	2819
DB		
2974	CCACGAATACTCGCTCAGGAGAAGGCACCTGACCACTACTCTCATCAGAAGCAACACTAA	3033
QY	CHCCTTCGAATTCCCTTC-ACCTCTCCAGAACCCAGAGCTTTGGCAGGGGAGACACCCGC	2878
DB		
3034	CCCCGACCAAT-CANCTCTTGACTTCAAAGGGGACCAGAGCTTTGGCAGGGGAGACACCCGC	3093
QY	CGCCCGCCGCCCGCAGAGSGTGGACTTCCAATCCAGGTGCACACCTCCCAGGCCATCTGGCCAGG	2938
DB		
3094	CGCCCGCCGCCCGCAGAGSGTGGACTTCCAATCCAGGTGCACACCTCCCAGGCCATCTGGCCAGG	3153
QY	CGGTGACTGTCTCGAGGCGAGCCGAGCTCAACGCCCTACAACCTCACTGCAAGAGTCGGGGC	2998
DB		
3154	CGGTGACTGTCTCGAGGCGAGCCGAGCTCAACGCCCTACAACCTCACTGCAAGAGTCGGGGC	3213
QY	TGAAGCGTACGCCCTTCGCTAAGCGCGGACGTAGTCCGCCCAAGACATCTTTGCTCCCTTT	3058
DB		
3214	TGAAGCGTACGCCCTTCGCTAAGCGCGGACGTAGTCCGCCCAAGACATCTTTGCTCCCTTT	3273
QY	CCACATTCATGAAGCCCAATGCGTGTACATAA	3093
DB		
3274	CCACATTCATGAAGCCCAATGCGTGTACATAA	3308
RESULT 3		
ABX71103		
ID	ABX71103 standard; cDNA: 4280 BP.	
XX	ABX71103;	
XX		
XX		
XX		
XX	05-MAR-2003 (first entry)	
XX		
XX	Novei human cDNA sequence #328.	
XX		
XX	Human: gene: ss: nervous system disorder; peripheral neuropathy;	
KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;	
KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;	
KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;	
KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;	
KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;	
KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;	
KW	coagulation disorder; cancer; tumour; inflammatory disease;	
KW	septic shock; Crohn's disease; anaphylaxis; proliferation; chemotactic;	
KW	differentiation; stem cell growth factor; haematopoiesis; chemokinetic;	
KW	haemostatic; antiinflammatory; expressed sequence tag; ESI.	
CS	Homo sapiens.	
PX	W020281731-A2.	
PN		
PD	17-OCT-2002.	
XX		
PF	29-JAN-2002; 2002WO-US01222.	
XX		
PR	30-JAN-2001; 2001US-0774528.	
XX		
PA	{HYSE-} HYSEQ INC.	
FA	(GOOD/) GOODRICH R W.	
XX		
PI	Tang TY, Jiao C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;	
PI	Xue A*, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;	
XX		
DR	WPI: 2003-058563/05.	
XX		
PT	Novel polypeptide useful for treating neurodegenerative diseases,	
PT	mucoid or lymphoid cell disorders, bone disorders, mechanical and	
PT	traumatic disorders, coagulation disorders, and inflammatory diseases	
PS	Claim 1: Page -; 612pp; English.	

Claim 1: Page -: 622pp; English.

XX

This invention relates to the cDNA sequences encoding an isolated novel human polypeptide. The protein encoded by the nucleic acid of the invention is useful for treating central and peripheral nervous system diseases (e.g. peripolar; neuropathy, Huntington's disease, amyotrophic lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease, Alzheimer's disease); autoimmune disease (e.g. systemic lupus erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus); myeloid or lymphoid cell disorders (e.g. anadmia and thrombocytopaenia); wounds, ulcers, burns; bone disorders (e.g. osteoporosis, osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head trauma); lung or liver fibrosis; reperfusion injury in various tissues; bacterial, viral or fungal infections; allergic conditions such as allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia); cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's disease, anaphylaxis). The protein may be used to inhibit the growth, infection or function of infectious agents such as bacteria, fungi, viruses, or to effect bodily characteristics, biorhythms or circadian cycles of rhythms. The protein may also have proliferation/differentiation, stem cell growth factor, hematopoiesis regulation, immune stimulating or suppressing, chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand and antiinflammatory activities. The cDNA sequences of the invention are useful for expressing recombinant protein for analysis. The present sequence represents a novel human cDNA sequence of the invention, this sequence is an expressed sequence tag (EST) and was identified using subtractive hybridization.

XX

SQ Sequence 4280 BP; 1185 A; 1165 C; 1002 G; 929 T; 0 other;

Query Match

Best Local Similarity 98.4%; Pred. No. 0;

Matches	Conservative	Mismatches	Indels	Gaps
Matches 3093;	Conservative 0	Mismatches 0	Indels 51;	Gaps 1;

QY	1	ATGAGGICAGAAGCCTTGGCTGCIA:ATTTCACACTGTACACATTTCCCTGGGGCTGGTTTC	68
Db	267	ATGAGGTGAGAAGCCTTGCTGCTATATTTCACATGTACACTTTGCTGGGGCTGGTTTC	326
QY	61	CACAAGAATCTCGAGCCCAATCAGTAITTCGCAITGGCAATATACAAACAGATCCGGTG	120
Db	327	CAGAAGATCTWAGCCCAATCAGTAITTCGCA:GGCAACTATACAAAACATATCCGGTG	385
QY	121	TTTGTGGGCCAACAGCCAGSGACGNAACACCACACAGAGGACACAGCTGGACATCCAGATG	189
Db	387	TTTGTGGGCCAACAGCCAGGACGGAACACCACACAGAGGACACAGCTGGACATCCAGATG	445
QY	181	ATTATGATCATGAACGGGAACCTCTNCATTCCTGCTAGGACCATATTATACITGTTGAT	249
Db	447	ATTATGATCATGAACGGGAACCTCTACATTCCTGCTAGGACCATATTATACITGTTGAT	506
QY	241	A:AGACATCACACACGCAAGAAATTAATGTAGCAAAAAAAGTCACATGGAAATCTAGA	300
Db	507	ATAAGACATCACACACGCAAGAAATTTATGTAGCAAAAAAAGTCACATGGAAATCTAGA	566
QY	301	CAGGCGGATGTAGACACATGCAGAAATGAAGGGAAAAACATTAAGGATGAGTGCACACACTTT	360
Db	567	CAGGCGGATGTAGACACATGCAGAAATGAAGGGAAAAACATTAAGGATGAGTGCACACACTTT	626
QY	361	AUTAAAGTTCTTAAAGAAACGATCATGCAATGTTGCT:GTGGAACTAATGCCCTTC	420
Db	627	ATTAAGTTCTTAAAGAAAGAGTAGTCATGTTGCTGTGGAACTAATGCCCTTC	686
QY	421	AACCCCTTCTTCAGAAACTATAAAGATGGATACATTGAACCATTCGGGATGAATTCAGC	480
Db	687	AACCCCTTCTTCAGAAACTATAAGATGSA:ACATGGNACCATTCGGGAGTAANTTCAGC	746
QY	481	GGAATGGCCAGATGCCCATATGATGCCAAACATGCCAACCTTGCAC TGTTTGCAGATGA	540
Db	747	GGAATGGCCAGATGCCCATATGATGCCAAACATGCCAACCTTGCAC TGTTTGCAGATGA	806
QY	541	AAACTATACTCAGCCACAGTACTGACTTCCTGCCCATTCAGCGAGTCATTTACCGGAGT	600
Db	807	AAACTATACTCAGCCACAGTACTGACTTCCTGCCCATTCAGCGAGTCATTTACCGGAGT	866

```
QY 1681 AATACAGATGGTCTGGGGAGCTGTACAAATTCCTTCTGGACACGA----- 1726
DB 1947 AATACAGATGGTCTGGGGAGCTGTACAAATTCCTTCTGGACACGAATGACATTCACAT 2006
QY 1727 -----ATGGGCATTCAGTTCCTCTTG 1749
DB 2007 CCTTACCAGATAATGAATGTCTTACACACAGTGTATGGGCATTCAGTTCCTCTTG 2066
QY 1750 CCCAGCACAAACACATCAGATTCGAGCGCTCAAGAGGGGTATGAGTCTAGGCGAGGAATG 1809
DB 2067 CCCAGCACAAACACATCAGATTCGAGCGCTCAAGAGGGGTATGAGTCTAGGCGAGGAATG 2126
QY 1810 CTGGACTGGAACATCTGCTGACTCAGCTGACACACAGACCTTTCGGGGCAGTGTCT 1869
DB 2127 CTGGACTGGAACATCTGCTGACTCAGCTGACACACAGACCTTTCGGGGCAGTGTCT 2186
QY 1870 TCCCATATACCAAGACAAAGAGAGTGTATTCGGGAAAGTACCTCAAAAGGCCACGAC 1929
DB 2187 TCCCATATACCAAGACAAAGAGAGTGTATTCGGGAAAGTACCTCAAAAGGCCACGAC 2246
QY 1930 CAGCTGGTTCGGCTCAGCCTCTCTGGCCATTCGAGTCATCTCGGGCTTCGTCATGGGGCC 1989
DB 2247 CAGCTGGTTCGGCTCAGCCTCTCTGGCCATTCGAGTCATCTCGGGCTTCGTCATGGGGCC 2306
QY 1990 GTCTTCTGGGCATCAGCTGTACTCGGTCTGTGATCATCGCGCAAGAGAGTGGCTGTG 2049
DB 2307 GTCTTCTGGGCATCAGCTGTACTCGGTCTGTGATCATCGCGCAAGAGAGTGGCTGTG 2366
QY 2050 GTGCAGCGCAAGAGAGAGTGTACCTACCTCGCGCGGCTCCATGAGCAGGTACCC 2109
DB 2367 GTGCAGCGCAAGAGAGAGTGTACCTACCTCGCGCGGCTCCATGAGCAGGTACCC 2426
QY 2110 AGCTCAGCGGCTCTTTTGGGACACTCATCTCAAAAGAGAGAGAGAGAGAGAGAGAGAG 2169
DB 2427 AGCTCAGCGGCTCTTTTGGGACACTCATCTCAAAAGAGAGAGAGAGAGAGAGAGAGAG 2486
QY 2170 AGCCACTCATGCACAAGCGCAAGTGTGCTGCTCGCGCAAGAGAGAGAGAGAGAGAGAG 2229
DB 2487 AGCCACTCATGCACAAGCGCAAGTGTGCTGCTCGCGCAAGAGAGAGAGAGAGAGAGAG 2546
QY 2230 AAGCAGACGACGACCTGTGACCTGTAGCGGCTTCCCGCAAGAGAGAGAGAGAGAGAGAG 2289
DB 2547 AAGCAGACGACGACCTGTGACCTGTAGCGGCTTCCCGCAAGAGAGAGAGAGAGAGAGAG 2606
QY 2290 CTGCAGCAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2349
DB 2607 CTGCAGCAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2666
QY 2350 ATCAATGCTGCACAAGAGACATGCCCCCGCATSGGCTCCCTGTGATTCGCCAGGACCTG 2409
DB 2667 ATCAATGCTGCACAAGAGACATGCCCCCGCATSGGCTCCCTGTGATTCGCCAGGACCTG 2726
QY 2410 CCGCTGCGGGCTTCCCGCAGACACATCCCGAGCGTGGTGGTCTCTGCCATCAGCGAGAG 2469
DB 2727 CCGCTGCGGGCTTCCCGCAGACACATCCCGAGCGTGGTGGTCTCTGCCATCAGCGAGAG 2786
QY 2470 GGCTACGACATGAGTACGTGACACAGCGCCAAATAGAGCGAGTGGCGCCACATGGCGCTG 2529
DB 2787 GGCTACGACATGAGTACGTGACACAGCGCCAAATAGAGCGAGTGGCGCGCCACATGGCGCTG 2846
QY 2530 GAGGACGCGCGCCACACTCTGAGTATAGACATCAAGAGAACATCTCAGCAGCAAGAGT 2589
DB 2847 GAGGACGCGCGCGCCACACTCTGAGTATAGACATCAAGAGAACATCTCAGCAGCAAGAGT 2906
QY 2590 CCAACCATGGGGTGGACCTTGTGAGAACCTGGACAGCTCCCGCCCAAGTTCACAG 2649
DB 2907 CCAACCATGGGGTGGACCTTGTGAGAACCTGGACAGCTCCCGCCCAAGTTCACAG 2966
QY 2650 CGGAGCGCTCCCTGGTCCCGGGAGCGCTCCCTGTCTCAGCAGCGGTCTAAGCAAGCGG 2709
DB 2967 CGGAGCGCTCCCTGGTCCCGGGAGCGCTCCCTGTCTCAGCAGCGGTCTAAGCAAGCGG 3026
QY 2710 CTGGAAATGCACCATCTCTCTTCTACGGGGTGTGACTATAGAGAGAGTACCCCAAGAAC 2769
```

```
DB 3027 CTGGAATGACACACTCCTCTCTACGGGGTGTGACTATAAGAGGAGCTACCCACGAAC 3086
QY 2770 TCCTCAGAGAAAGCCAGGCGCACCCTCTCAAAAGAAACAACTAACTCCTCCTCAAT 2829
DB 3087 TCCTCAGAGAAAGCCAGGCGCACCCTCTCAAAAGAAACAACTAACTCCTCCTCAAT 3146
QY 2830 TCCTCAGCTCTCCAGAAACAGAGCTTTGGCAGGGAGAGAACCCGCGCGCCGCCG 2889
DB 3147 TCCTCAGCTCTCCAGAAACAGAGCTTTGGCAGGGAGAGAACCCGCGCGCCGCCG 3206
QY 2890 CAGAGGTGTGACTCCATCCAGGTGCACAGCTCCAGCGCTCTGGCCAGCGCTGACTGTC 2949
DB 3207 CAGAGGTGTGACTCCATCCAGGTGCACAGCTCCAGCGCTCTGGCCAGCGCTGACTGTC 3266
QY 2950 TCAGGCGAGCCCGAGCGCTCAACCGCTCAACTCACTGACAAGGTGCGGGGCTGAAGCGTAG 3009
DB 3267 TCAGGCGAGCCCGAGCGCTCAACCGCTCAACTCACTGACAAGGTGCGGGGCTGAAGCGTAG 3326
QY 3010 CCCTCGCTAAAGCGGAGCGTACCCCGCAACCATCTTTCCTCCCTTTCCACATCATG 3069
DB 3327 CCCTCGCTAAAGCGGAGCGTACCCCGCAACCATCTTTCCTCCCTTTCCACATCATG 3386
QY 3070 AAGCCCAATGATCGTGTACATAA 3093
DB 3387 AAGCCCAATGATCGTGTACATAA 3410
RESULT 4
AAH18729
ID AAH18729 Standard; cDNA; 6050 BP.
XX AAH18729;
AC AAH18729;
DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:19004.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
XX 26-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai I, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX Claim 8; SEQ ID 19004; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
```

CC oligonucleotide comprises at least 15 nucleotides; or (2) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03156 to AAH3628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB35893 represent human amino acid sequences; and AAH3629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
50 Sequence 5050 BP; 1639 A; 1525 C; 1363 G; 1532 T; 0 other;  
Query Match 96.0%; Score 2969.6; DB 22; Length 6060;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2994; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 96 CAACATATACAAACAGATCGCGTGTGTGGGCCAACAGCCAGCAGCAACCCACACA 155  
DB 1 CAACATATACAAACAGATATCGCGTGTGTGGGCCAACAGCCAGCAGCAACCCACACA 60  
QY 156 GAGGCACAGGCTGCATCCAGATGATATGATCATGATCAAGGAAACCTCTACATGCTGC 215  
DB 61 GAGGCACAGGCTGCATCCAGATGATATGATCATGATCAAGGAAACCTCTACATGCTGC 120  
QY 216 TAGGACCATATTTATCTGTGTATATAGACATACACAGCAAGCAAAATTTATTCTAG 275  
DB 12: TAGGACCATATTTATCTGTGTATATAGACATACACAGCAAGCAAAATTTATTCTAG 180  
QY 276 CAAAAAGCTGCATGGAATCTAGACAGGCGCATGTAGACACATGCGAGATGAAGGAAA 335  
DB 181 CAAAAAGCTGCATGGAATCTAGACAGGCGCATGTAGACACATGCGAGATGAAGGAAA 240  
QY 336 ACATAGATGATGCGCAACACTTTATTAAGTCTCTTAAGAAACCAACGATGATGATT 395  
DB 241 ACATAGATGATGCGCAACACTTTATTAAGTCTCTTAAGAAACCAACGATGATGATT 300  
QY 396 GTTGTCTGTGGAACTAATGCTTCAAGCCCTCTGTCAGAACTATAAGATGGAATGATT 455  
DB 301 GTTGTCTGTGGAACTAATGCTTCAAGCCCTCTGTCAGAACTATAAGATGGAATGATT 350  
QY 456 GAAACCATTCGGGATGAATTCAGCGGATGCGCAGATGCGCATATGATGCCAACAATCC 515  
DB 361 GAAACCATTCGGGATGAATTCAGCGGATGCGCAGATGCGCATATGATGCCAACAATCC 420  
QY 516 CAACGTTCGACTGTTCGAGATGGAAGAACTATCTCAGCTACAGTGAATGCTTCCTGTC 575  
DB 421 CAACGTTCGACTGTTCGAGATGGAAGAACTATCTCAGCTACAGTGAATGCTTCCTGTC 480  
QY 576 CATTCAGCGAGTCATTTACCGAGTCTTGGAGAAAGCCCTACCTGCGGACCGTCACAGCA 635  
DB 481 CATTCAGCGAGTCATTTACCGAGTCTTGGAGAAAGCCCTACCTGCGGACCGTCACAGCA 540  
QY 636 CCATTCAAAATGGTTGAAGAACCATATCTTCTCAAGCCGCGGATACCGGAGATTATAT 695  
DB 541 CCATTCAAAATGGTTGAAGAACCATATCTTCTCAAGCCGCGGATACCGGAGATTATAT 600  
QY 696 CTACTCTTCTTCAGGGAATATAGAGTGGAGTATTAACACCATGGGAAAGTAGTCTTCCC 755  
DB 601 CTACTCTTCTTCAGGGAATATAGAGTGGAGTATTAACACCATGGGAAAGTAGTCTTCCC 660  
QY 756 AAGATGCTCAGGTTGTAGCAATGATATGGAGGATCTCAAGAGTCTCTGGAGAAACA 815  
DB 661 AAGATGCTCAGGTTGTAGCAATGATATGGAGGATCTCAAGAGTCTCTGGAGAAACA 720  
QY 816 GTGAGGCTGCTCCCTGAAGGCGCGCTTCAACCTGCTCAGTCTCCGAGACTCTCATTTTA 875

DB 721 GTGACGCTGCTTCTGAGGCGGCTTAACTGCTCAGTTCCTGAGACTCTCATTTTA 780  
QY 876 TTCAACATTTCTCAGGCAATACAGATGTGATTCCTATCAAGGGGCGTGAITGTTCCT 935  
DB 751 TTCAACATTTCTCAGGCAATACAGATGTGATTCCTATCAAGGGGCGTGAITGTTCCT 840  
QY 936 GCGAACGTTTCTCAGCTTATACAGATCCCTGGGTCTGCGAGTCTGTCCTTACAT 995  
DB 841 GCGAACGTTTCTCAGCTTATACAGATCCCTGGGTCTGCGAGTCTGTCCTTACAT 900  
QY 996 GCTTGACATTCGACATGCTTTTACTGGAGATTTCAAGGAACAGAAATCTCTGATTCCAC 1055  
DB 901 GCTTGACATTCGACATGCTTTTACTGGAGATTTCAAGGAACAGAAATCTCTGATTCCAC 960  
QY 1056 CTGGACACAGGTTCTGATGAACAGATTCCTAAGCCAGGCGAGGTTGCTGTGCTGCTC 1115  
DB 961 CTGGACACAGGTTCTGATGAACAGATTCCTAAGCCAGGCGAGGTTGCTGTGCTGCTC 1020  
QY 1115 AACTCTCTTACAAAGATATGCAACTCCAAATGCTTCCCTGATGATACCTGAACTTCA 1175  
DB 1021 ATCTCTCTTACAAAGATATGCAACTCCAAATGCTTCCCTGATGATACCTGAACTTCA 1080  
QY 1176 CAAGACGACCGGCTCATGATGAGGAGTSCCTCTCCATCTTCAACAGGCGCATGGTTCCT 1235  
DB 1081 CAAGACGACCGGCTCATGATGAGGAGTSCCTCTCCATCTTCAACAGGCGCATGGTTCCT 1140  
QY 1236 GAGAACATGATCAGATACCGGCTTACCAAAATGCAATGCAAGTGCAGACACTGCTGGCATA 1295  
DB 1141 GAGAACATGATCAGATACCGGCTTACCAAAATGCAATGCAAGTGCAGACACTGCTGGCATA 1200  
QY 1296 TCAGAAATCAGACTGCTGTTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTTCG 1355  
DB 1201 TCAGAAATCAGACTGCTGTTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTTCG 1260  
QY 1356 CAGAAATAGAAATAGTGGTTTCTTAATGACAGGCTTTTCTCTGAGGAGATGAGTGTTA 1415  
DB 1261 CAGAAATAGAAATAGTGGTTTCTTAATGACAGGCTTTTCTCTGAGGAGATGAGTGTTA 1320  
QY 1416 CAACCTCTG-AAAAATGCACTATGATGAGTGAAGCAAAAGGATCATGGCATGCA 1474  
DB 1321 CAACCTCTGCAAAATGCACTATGATGAGTGAAGCAAAAGGATCATGGCATGCA 1380  
QY 1475 TGACAGACAGCAGCTCTCTGTATGTGGGTCTCTTACCTGTGTATAAAGTTTCCCC 1534  
DB 1381 TGACAGACAGCAGCTCTCTGTATGTGGGTCTCTTACCTGTGTATAAAGTTTCCCC 1440  
QY 1535 CTGGCGGTGTGAACCACTGAGAGTCT-AAAAAACCCTGATTGCTCCAGAGACCA 1593  
DB 1441 CTGGCGGTGTGAACCACTGAGAGTCT-AAAAAACCCTGATTGCTCCAGAGACCA 1500  
QY 1594 TATTGTGATGATAAAGAAAGTGTGCTGCTGAGGCAATTTATCACCACAGCAGACTG 1653  
DB 1501 TATTGTGATGATAAAGAAAGTGTGCTGCTGAGGCAATTTATCACCACAGCAGACTG 1560  
QY 1654 ACTTTGACGAGCATAGAGGTTGGCAATACAGATGCTCTGGGGGACTGTCACAATTC 1713  
DB 1561 ACTTTGACGAGCATAGAGGTTGGCAATACAGATGCTCTGGGGGACTGTCACAATTC 1620  
QY 1714 TTTGTGGCACTGAATGGGCAATTCAGTTCCCTCTTGGCCAGCAACACACATCAGATTG 1773  
DB 1621 TTTGTGGCACTGAATGGGCAATTCAGTTCCCTCTTGGCCAGCAACACACATCAGATTG 1680  
QY 1774 AGGCTCAGAGAGGTTATGATCTAGGGGAGGATGCTGGACTGGAGCATCTCTCTTGAC 1833  
DB 1681 AGGCTCAGAGAGGTTATGATCTAGGGGAGGATGCTGGACTGGAGCATCTCTCTTGAC 1740  
QY 1834 TCACCTGACGACACAGACCTTTTGGGGCAGTGTCTTCCCATATATCAGAGACAAG 1893  
DB 1741 TCACCTGACGACACAGACCTTTTGGGGCAGTGTCTTCCCATATATCAGAGACAAG 1850  
QY 1894 GGAGTGAATTCGGGAAGTTACTTCAAGGCGCAGCAGCTGGTTCGCTCAGCTCTTG 1953

Db 1801 GGAGTGAATTCGGGAAGTTACCTCAAGGCCACGACAGCTGGTTCCCGTCACCTCTTG 1860  
Qy 1954 GGCATTGCAAGTCACTGTCGCTTCGTCATGGGGCGCTCTCTCTCGGGCATCACCGTCTAC 2013  
Db 1861 GCCATTGCAAGTCACTGTCGCTTCGTCATGGGGCGCTCTCTCTCGGGCATCACCGTCTAC 1920  
Qy 2014 TCGCTGTGATCATCGGGCGCAAGACAGTGGCTGTGGTGCAGCGCAAGAGAGAGTCTC 2073  
Db 1921 TCGCTGTGATCATCGGGCGCAAGACAGTGGCTGTGGTGCAGCGCAAGAGAGAGTCTC 1980  
Qy 2074 ACCACTCCGGCGGGCTTCATGACGACGCTACCAAGCTCAGCGGCTCTTTGGGAC 2133  
Db 1981 ACCACTCCGGCGGGCTTCATGACGACGCTACCAAGCTCAGCGGCTCTTTGGGAC 2040  
Qy 2134 ACTCAATCAAGACCAAGACGGAGGCCATCTCTCAAGGCCACTCATGCAACAGCGCAAG 2193  
Db 2041 ACTCAATCAAGACCAAGACGGAGGCCATCTCTCAAGGCCACTCATGCAACAGCGCAAG 2100  
Qy 2194 CTGCGCACTCCGGCAACAGCGGAGATGCTCATTAAGCAGACGACGACCTGGAC 2253  
Db 2101 CTGCGCACTCCGGCAACAGCGGAGATGCTCATTAAGCAGACGACGACCTGGAC 2160  
Qy 2254 CTGACGGCTCTCCGACCGACAGTCAAGCCCAAGCGCTGACGAGAGCGGAAGGCCAC 2313  
Db 2161 CTGACGGCTCTCCGACCGACAGTCAAGCCCAAGCGCTGACGAGAGCGGAAGGCCAC 2220  
Qy 2314 CGGGCAGCGGAGTGGAGAGGAACAGACCTCATATGCTGCTGACAAAGACATG 2373  
Db 2221 CGGGCAGCGGAGTGGAGAGGAACAGACCTCATATGCTGCTGACAAAGACATG 2280  
Qy 2374 CGCCCATGGCTCGCTCTGATTCGACGAGAGCTGCGCTCGCGCTCGCCAGCGAC 2433  
Db 2281 CGCCCATGGCTCGCTCTGATTCGACGAGAGCTGCGCTCGCGCTCGCCAGCGAC 2340  
Qy 2434 ATCCGAGCGTGGTGTCTTGCCATACGACAGAGGCTACGACATGASTACCTGGAC 2493  
Db 2341 ATCCGAGCGTGGTGTCTTGCCATACGACAGAGGCTACGACATGASTACCTGGAC 2400  
Qy 2494 CAGCCCAATCAGCAGCTGGCCAGATGGCGTGGAGGACGCGCGCCACACCTGGAG 2553  
Db 2401 CAGCCCAATCAGCAGCTGGCCAGATGGCGTGGAGGACGCGCGCCACACCTGGAG 2460  
Qy 2554 TATAAGACCATCAAGGAACATCTCAGCAGACAGATGCCAATCATGGGTGAACCTTGTG 2613  
Db 2461 TATAAGACCATCAAGGAACATCTCAGCAGACAGATGCCAATCATGGGTGAACCTTGTG 2520  
Qy 2614 GAGACCTGGACAGCTGCGCCGCAAGTTCACACAGGGAGGCGCTCCCTGGGTCGCCG 2673  
Db 2521 GAGACCTGGACAGCTGCGCCGCAAGTTCACACAGGGAGGCGCTCCCTGGGTCGCCG 2580  
Qy 2674 GGAGCTGCGCTGTCTCAGACCGGTCTAAGCAAGCGGCTGGAATGACACCTCTCTCC 2733  
Db 2581 GGAGCTGCGCTGTCTCAGACCGGTCTAAGCAAGCGGCTGGAATGACACCTCTCTCC 2640  
Qy 2734 TAGCGGTGTACTATTAAGAGGAGCTACCCGACGAAGTCTCAGAGAGCCACGAGCC 2793  
Db 2641 TAGCGGTGTACTATTAAGAGGAGCTACCCGACGAAGTCTCAGAGAGCCACGAGCC 2700  
Qy 2794 ACCACTCTCAAGAAACAACTACTCTCTCAATCTCTCACTCTCTCCAGAAACAG 2853  
Db 2701 ACCACTCTCAAGAAACAACTACTCTCTCAATCTCTCACTCTCTCCAGAAACAG 2760  
Qy 2854 AGCTTGGAGGAGACAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2913  
Db 2761 AGCTTGGAGGAGACAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2820  
Qy 2914 CACAGTCCAGGCTCTGCGAGGCGGTGACTGTCTGAGGAGCGCCACGCTCAAGCCG 2973  
Db 2821 CACAGTCCAGGCTCTGCGAGGCGGTGACTGTCTGAGGAGCGCCACGCTCAAGCCG 2880  
Qy 2974 TACAACCTCACTGACAAAGTCTGGGGCTGAAGGCTGACGCGCTCGCTAAAGCGCGACGCTAC 3033  
Db 2881 TACAACCTCACTGACAAAGTCTGGGGCTGAAGGCTGACGCGCTCGCTAAAGCGCGACGCTAC 2940

Qy 3034 CCCAAACCACTCTTGTCCCTTCCACATCCATGAAGCCCAATGAIGCTGTACATAA 3093  
Db 2941 CCCAAACCACTCTTGTCCCTTCCACATCCATGAAGCCCAATGAIGCTGTACATAA 3000

## RESULT 5

AAV44295  
ID AAV44295 standard; cDNA; 3550 BP.

AC AAV44295;

XX 06-OCT-1998 (first entry)

XX Human: secreted protein clone CJ145\_1 cDNA.

XX Secreted protein; human fetal brain; nutrition; cytokine; stimulant;  
cell proliferation; differentiation; immune system; suppressor; ligand;  
KW regulator; hematopoiesis; tissue growth; activin; inhibitor; haemostatic;  
KW chemotaxis; chemokine; thrombosis; receptor; cadherin; tumour;  
KW anti-inflammatory; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
CDS :26..3053

FT /\*tag- a  
FT /product= secreted protein  
FT /note= \*isolated from clone CJ145\_1\*

XX WO9827205-A2.

XX 25-JUN-1998.

XX 17-DEC-1997; 97WO-US23330.

XX 16-DEC-1997; 97US-0951872.

XX 18-DEC-1996; 96US-0769192.

XX 13-JAN-1997; 97US-0783401.

XX (GEWY ) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavalie ER, McCoy JM, Merberg D;  
Racie LA, Spaulding V, Treacy M;

XX WPI: 1998-362774/31.

XX P-PSDB: AAW64221.

XX New polynucleotides and secreted proteins - obtained from human  
fetal brain, human adult testes, human adult brain and human adult  
salivary gland cDNA libraries

XX Claim 17a; Page 69-71; 10pp; English.

XX This sequence encodes a novel secreted protein from clone CJ145\_1  
isolated from a human fetal brain cDNA library. This protein has  
applications for nutritional use, cytokine and cell  
proliferation/differentiation activity, immune stimulating or  
suppressing activity, hematopoiesis regulating activity, tissue growth  
activity, activin/inhibin activity, chemotactic/chemokinetic activity,  
haemostatic and thrombotic activity, receptor/ligand activity,  
anti-inflammatory activity, cadherin/tumour invasion suppressor activity,  
tumour inhibition activity and other activities.

XX Sequence 3550 BP; 957 A; 994 C; 856 G; 742 T; 1 other;

Query Match 88.8%; Score 2746.2; DB 19; Length 3550;

Best Local Similarity 94.5%; Pred. No. 0;

Matches 2923; Conservative 1; Mismatches 4; Indels 165; Gaps 1;

Qy 1 ATGAGGTCAGAGGCTTGGCTGCTATATTCACACTGCTACACTTTCGCGGCTGTTTC 60

Db 126 ATGAGGTCAGAGGCTTGGCTGCTATATTCACACTGCTACACTTTCGCGGCTGTTTC 185



QY 61 CCAGAGANTCTGAGCCAAATCAGTATTGCGAATGGCAACTATACAAACAGTAGTCCGGTG 120  
DB 186 CCAGAGANTCTGAGCCAAATCAGTATTGCGAATGGCAACTATACAAACAGTAGTCCGGTG 245  
QY 121 TCTGGGCCACAAGCCAGAGCAGCAACAGCACACAGAGCCAGAGCGGTGGACATCCAGATG 180  
DB 246 TTTGTGGGCCACAAGCCAGAGCAGCAACAGCACACAGAGCCAGAGCGGTGGACATCCAGATG 325  
QY 181 ATTAATGATCATGAAGGAGACCTGCTGATTTGCTGCTAGGAGCAATATTTATATGTTGAT 240  
DB 306 ATTAATGATCATGAAGGAGACCTGCTGATTTGCTGCTAGGAGCAATATTTATATGTTGAT 365  
QY 241 ATAGACATCATACACAGGAGAGAAATTTATTTGAGCAAAAACCTGACATGAAATGTAGA 300  
DB 366 ATAGACATCATACACAGGAGAGAAATTTATTTGAGCAAAAACCTGACATGAAATGTAGA 425  
QY 301 CAGGCCATGTAGACACATGAGAAATGAAGGAGAAACATAGAGATGAGTCCCAAACTTT 360  
DB 426 CAGGCCATGTAGACACATGAGAAATGAAGGAGAAACATAGAGATGAGTCCCAAACTTT 485  
QY 361 ATTAAGTTCTTAAAGAAATGAGGATGATGATTTGCTGCTGAGAACTATAGCTTC 420  
DB 486 ATTAAGTTCTTAAAGAAATGAGGATGATGATTTGCTGCTGAGAACTATAGCTTC 545  
QY 421 AACCTTCTCCAGAACTATAAGATGATACATATGGAACCATTCGGGGAGTAATTCAGC 480  
DB 546 AACCTTCTCCAGAACTATAAGATGATACATATGGAACCATTCGGGGAGTAATTCAGC 605  
QY 481 GGAATGGCCAGATGCCCATATGATGCCAAACATGCCAACTTGGACPTTTGAGATGGA 540  
DB 606 GGAATGGCCAGATGCCCATATGATGCCAAACATGCCAACTTGGACPTTTGAGATGGA 665  
QY 541 AAATTAATCTCAGCACAGTACGATCTCTGCGAATTCAGCGAGTCATTTACCGGAGT 600  
DB 666 AAATTAATCTCAGCACAGTACGATCTCTGCGAATTCAGCGAGTCATTTACCGGAGT 725  
QY 601 GTTGAGAAACCTTACCTCGGACCGTCAAGCAGCATTCANAAATGCTTGAAGAAGACCA 660  
DB 726 GTTGAGAAACCTTACCTCGGACCGTCAAGCAGCATTCANAAATGCTTGAAGAAGACCA 785  
QY 661 TACTTTGTTCAGCGGTGGATACGGAGATATATCTTACTTCTTCAGGGAATAGCA 720  
DB 786 TACTTTGTTCAGCGGTGGATACGGAGATATATCTTACTTCTTCAGGGAATAGCA 845  
QY 721 GTTGAGATATAACACATGGGAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAGAAAT 780  
DB 846 GTTGAGATATAACACATGGGAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAGAAAT 905  
QY 781 GNTATGGAGAGATCTCAAGAGTCTGAGAAACAGTGGAGTGGTTCTTGAAGCGCGC 840  
DB 906 GAATGGAGAGATCTCAAGAGTCTGAGAAACAGTGGAGTGGTTCTTGAAGCGCGC 965  
QY 841 TTGAAGTCTCAGTTCTCGAGAGTCTCATTTTATTTCAACATCTCCAGCGAGTTACA 900  
DB 966 TTGAAGTCTCAGTTCTCGAGAGTCTCATTTTATTTCAACATCTCCAGCGAGTTACA 1025  
QY 901 GATGTGATGTATACAGCGCGGTGATGTTGCTGCTGGCAAGGTTTCTACACCTTATAAC 960  
DB 1026 GATGTGATGTATACAGCGCGGTGATGTTGCTGCTGGCAAGGTTTCTACACCTTATAAC 1085  
QY 961 ACATCTCCTGGTGGAGTCTGCTGCTAAGCATGCTTGACATGCTTGACATGCTGCTTACT 1020  
DB 1086 ACATCTCCTGGTGGAGTCTGCTGCTAAGCATGCTTGACATGCTTGACATGCTGCTTACT 1145  
QY 1021 GGGAGATTCAGGAACAGAACTCTCTGATTCACCTGGACACACAGTTCCTGATGAACGA 1080  
DB 1146 GGGAGATTCAGGAACAGAACTCTCTGATTCACCTGGACACACAGTTCCTGATGAACGA 1205  
QY 1081 GTTCTTAAGCCAGCGCAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 1206 GTTCTTAAGCCAGCGCAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265

QY 1141 TCCATAGAGTTCCTCTGATGATACCTGAACCTCATCAAGACGACCCGCTCATGATGAG 1200  
DB 1266 TCCATAGAGTTCCTCTGATGATACCTGAACCTCATCAAGACGACCCGCTCATGATGAG 1325  
QY 1201 GCAGTGGCCTCCATCTTCAACAGAGCCATGGTTCCTGAGAACTATGCTCAGATACCGCTT 1260  
DB 1326 GCAGTGGCCTCCATCTTCAACAGAGCCATGGTTCCTGAGAACTATGCTCAGATACCGCTT 1385  
QY 1261 ACCAAATATGAGTGGACACAGCTGCTGGGCCATATCAGAAATCAGATGCTGGTTTCTG 1320  
DB 1386 ACCAAATATGAGTGGACACAGCTGCTGGGCCATATCAGAAATCAGATGCTGGTTTCTG 1445  
QY 1321 GATCAGAGAGAGGAACTATCTTGAAGTTTGTGCGCAGAAATAGTGGTTTCTTA 1380  
DB 1446 GATCAGAGAGAGGAACTATCTTGAAGTTTGTGCGCAGAAATAGTGGTTTCTTA 1505  
QY 1381 AATGACAGCCTTTTCTGAGAGAGATGATGTTTACAACCTGAAAAATGCAGTATGAT 1440  
DB 1506 AATGACAGCCTTTTCTGAGAGAGATGATGTTTACAACCTGAAAAATGCAGTATGAT 1565  
QY 1441 GAGTTCGAGAGACAAAGGATCATGGGCATGACAGCTGGACAGAGCAAGCAGTCTCTGTAT 1500  
DB 1566 GAGTTCGAGAGACAAAGGATCATGGGCATGACAGCTGGACAGAGCAAGCAGTCTCTGTAT 1625  
QY 1501 GTTGGCTTCTTACCTGTGTATTAAGCTTCCCTTGGCCGCTGTGAACGACATGGGAAG 1560  
DB 1526 GTTGGCTTCTTACCTGTGTATTAAGCTTCCCTTGGCCGCTGTGAACGACATGGGAAG 1685  
QY 1561 TGTAAAAAAGCTGTTATGCTCCAGAGACCCATTTGTTGGATGATTAAGGAAGTGGT 1620  
DB 1686 TGTAAAAAAGCTGTTATGCTCCAGAGACCCATTTGTTGGATGATTAAGGAAGTGGT 1745  
QY 1621 GCCTCAGCAGCTTTATCACCACCAACAGAGCTGACTTTTIGAGCAGGACATAGAGCGTGGC 1680  
DB 1746 GCCTCAGCAGCTTTATCACCACCAACAGAGCTGACTTTTIGAGCAGGACATAGAGCGTGGC 1805  
QY 1681 AATACAGATGGTCTGGGAGCTGTCCAAATCTCTTGTGGCAGCTGAATGGCAGTATCCAGT 1740  
DB 1806 AATACAGATGGTCTGGGAGCTGTCCAAATCTCTTGTGGCAGCTGAATGGCAGTATCCAGT 1853  
QY 1741 TCCCTCTTGGCCAGCACACACATCAGATTCGAGCGCTCAAGAGGGTATGATCTAGG 1800  
DB 1854 TCCCTCTTGGCCAGCACACACATCAGATTCGAGCGCTCAAGAGGGTATGATCTAGG 1853  
QY 1801 GGAGGAATCTGGACTGGAAGCATCTGCTTGACTACCTGACACAGACAGACCCCTTTGGG 1860  
DB 1854 GGAGGAATCTGGACTGGAAGCATCTGCTTGACTACCTGACACAGACAGACCCCTTTGGG 1853  
QY 1861 GCAGTCTCTTCCCATATACCAACAGAGAGGAGTGTTCGGGAAAGTTACTCTAAA 1920  
DB 1854 GCAGTCTCTTCCCATATACCAACAGAGAGGAGTGTTCGGGAAAGTTACTCTAAA 1880  
QY 1921 GGCCAGCAGCAGCTGTTTCCCTCACCCTCTTGGCCATTCGAGTCACTCGCTGGCTTCTGTC 1980  
DB 1881 GGCCAGCAGCAGCTGTTTCCCTCACCCTCTTGGCCATTCGAGTCACTCGCTGGCTTCTGTC 1940  
QY 1981 ATGGGGCGCTCTTCTGGGCACTACCGTCTACTGCTGCTGTGATCATCGCGGCAAGAC 2040  
DB 1941 ATGGGGCGCTCTTCTGGGCACTACCGTCTACTGCTGCTGTGATCATCGCGGCAAGAC 2000  
QY 2041 GTGGGTGTGGTCCAGCGCAAGGAGAGGAGTCAACCCACTCGCCCGGGGCTCCATGAGC 2100  
DB 2001 GTGGGTGTGGTCCAGCGCAAGGAGAGGAGTCAACCCACTCGCCCGGGGCTCCATGAGC 2060  
QY 2101 AGCGTCACAAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGGAG 2160  
DB 2061 AGCGTCACAAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGGAG 2120  
QY 2161 GGCATCTCTACGCGCTATSCACAGCGCAGCTCGCCACTCCCGCAACACCGCCAG 2220  
DB 2121 GGCATCTCTACGCGCTATSCACAGCGCAGCTCGCCACTCCCGCAACACCGCCAG 2180  
QY 2221 ATGCTCATTAAGAGAGACCCAGCACCCACTGGAGCTGACCGGCTCCCGCACCCAGAGTCA 2280



[illegible]

RESULT 6  
AAF98469  
ID AAF98469 standard; cDNA; 3550 BP.

AC AAF98469;

DT 07-JUN-2001 (first entry)

Human cDNA clone CJ145; sequence SEQ ID i60.

Human; secreted protein; nutrient; cytokine modulator; proliferation;  
KW differentiation; immune system modulator; tissue growth; chemotactic;  
KW haemostatic; thrombolytic; anti-inflammatory; tumor inhibition; ss;

KW	haematopoiesis.
XX	
OS	Homo sapiens.
XX	
PN	WO200119988-A1.
XX	
PD	22-MAR-2001.
XX	
XX	14-SEP-2000; 2000WO-US25135.
XX	
XX	17-SEP-1999; 99US-0398829.
PR	
XX	(GENY ) GENETICS INST INC.
PA	
XX	
PI	Jacobs K, McCoy JM, LaVallic ER, Collins-Racie LA, Evans C;
PI	Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;
XX	
DR	WPI: 2001-244801/25.
DR	P-FSGB; AAB90731.
XX	
XX	Isolated nucleic acids encoding polypeptides, useful for modulating
PI	e.g. cytokine and cell proliferation/differentiation activity, the
PI	immune system and hematopoiesis regulating activity -
XX	
PS	Disclosure; Page 486-487; 557pp; English.
XX	
XX	Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
CC	proteins AAB96667 - AAB90750. The cDNA clones are isolated from various
CC	tissue types, and may be used in the prevention, treatment and diagnosis
CC	of diseases associated with inappropriate protein expression. The
CC	polypeptides and nucleic acids may be used as nutrients or to modulate
CC	cytokine and cell proliferation/differentiation activity and may also be
CC	involved in modulation of the immune system. The cDNA sequences,
CC	proteins, their agonists and/or antagonists exhibit haematopoiesis
CC	regulating activity; tissue growth activity; activin/inhibin activity;
CC	chemotactic/chemokinetic activity; haemostatic and thrombolytic
CC	activity; receptor/ligand activity; anti-inflammatory activity;
CC	haematopoiesis activity; cadherin/tumour suppressor activity; and/or
CC	tumour inhibition activity. Included in the invention are probes
CC	represented in AAF98490 - AAF98572 which are specific for the cDNA clones
CC	encoding the secreted proteins.

Sequence 3550 BP; 957 A; 994 C; 856 G; 742 T; 1 other; 2250

```
Query Match      88.8%; Score 2746.2; DB 22; Length 3550;
Best local Similarity 94.5%; Pred. No. 0;
Matches 2923; Conservative 1; Mismatches 4; Indels 165; Gaps 1
```

[illegible]







QY 264: GTTCACAGCGGAGCGCTCTCTGGTCCCGGAGAGTCCCTCTCTCAGACGGCTCTA 2700  
 DB 2689 GTTCCACAGCGGAGCGCTCTCTGGTCCCGGAGAGTCCCTCTCTCAGACGGCTCTA 2748  
 QY 2701 AGCAAGCGCTGGAATGACACCACTCTCTTCTTCTACGSGGTTGACTATAGAGAGCTAC 2750  
 DB 2749 AGCAAGCGCTGGAATGACACCACTCTCTTCTTCTACGSGGTTGACTATAGAGAGCTAC 2808  
 QY 2751 CCCAGAACTCCCTCAGAGAGCCACAGGCGACC-ACCTCAAGAAACACACACTAA 2819  
 DB 2809 CCCAGAACTCCCTCAGAGAGCCACCTGACCACTTCTCAACAGAGCAACACTAA 2868  
 QY 2820 CTCCTCCAACTCTCTC-ACCTCTCCAGAAACCAAGAGCTTTGGCAGGCGAGCAACCCGC 2878  
 DB 2859 CCCAGAACTCCCTCAGAGAGCCACAGGCGACCAGAGCTTTGGCAGGCGAGCAACCCGC 2928  
 QY 2879 CCGCCGCGCCGAGAGGCTGAGACTCCATCCAGGTCAGAGGCGAGAGCTTTGGCAGGCGAG 2938  
 DB 2929 CCGCCGCGCCGAGAGGCTGAGACTCCATCCAGGTCAGAGGCGAGAGCTTTGGCAGGCGAG 2989  
 QY 2939 CCGTCACTCTCTCGAGGAGCCAGGCTCAAGCGCTACAACTTCACTGACAAAGGTGAGGGC 2998  
 DB 2989 CCGTCACTCTCTCGAGGAGCCAGGCTCAAGCGCTACAACTTCACTGACAAAGGTGAGGGC 3048  
 QY 2999 TGAAGGTTAGCGCTCGCTTAAGGCGGAGCTTACCGCCCAACCACTCTTTGCTCCCTTT 3056  
 DB 3049 TGAAGGTTAGCGCTCGCTTAAGGCGGAGCTTACCGCCCAACCACTCTTTGCTCCCTTT 3108  
 QY 3059 CCACATCCATGAAGCCCAATGATGCTGTATATAA 3093  
 DB 3109 CCACATCCATGAAGCCCAATGATGCTGTATATAA 3143

## RESULT 8

AAS68807  
 ID AAS68807 standard; cDNA; 3039 bp.  
 AC AAS68807;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #4611.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200175067-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-0508531.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0549167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YI;  
 XX MF: 2001-639362/73.  
 DR P-PSDB: ABG04620.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 1; SEQ ID No 4611; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (i) and

CC polypeptide (ii) sequences. (i) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (ii). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (i) is useful in gene therapy techniques  
 CC to restore normal activity of (ii) or to treat disease states involving  
 CC (ii). (ii) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (ii) and its binding partners are useful in medical  
 CC imaging of sites expressing (ii). (i) and (ii) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

Sequence 3039 bp; 741 A; 869 C; 781 G; 648 T; 0 other;

Query Match 71.8%; Score 2221.8; DB 23; Length 3039;  
 Best Local Similarity 91.2%; Pred. No. 0;  
 Matches 248; Conservative 0; Mismatches 167; Indels 65; Gaps 2;  
 QY 519 CTTTCACCTGTTTGCAGATGGAAGAACTATATCTACGACAGTACTGCTTCCCAT 578  
 DB 359 CTTGACTCTGCTCCAGATGGAAGAACTATATCTACGACAGTACTGCTTCCCAT 458  
 QY 579 TCACGACGTCATTTACGCGAGCTCTGGGAGAGCCCTACCTGGGACCGTCAAGCAGA 638  
 DE 459 TCACGACGTCATTTACGCGAGCTCTGGGAGAGCCCTACCTGGGACCGTCAAGCAGA 518  
 QY 639 TTCAAAATGTTTGAAGAAACCACTTTGTTTCAACGCGTGGATTACGAGATTATCTA 698  
 DE 519 TTCAAAATGCTTGAAAGAAACCACTTTGTTTCAACGCGTGGATTACGAGATTATCTA 578  
 QY 699 CTCTCTCTTCAGGGAATAGCAGTGGAGATATACACCATGCGGAGAGTAGTTTCCCAAG 758  
 DE 579 CTCTCTCTTCAGGGAATAGCAGTGGAGATATACACCATGCGGAGAGTAGTTTCCCAAG 638  
 QY 759 AGTGCTCAGGTTTGTAAAGATGATATCGGAGGATCTCAAGAGTCTCTGAGAAACAGTG 818  
 DE 639 AGTGCTCAGGTTTGTAAAGATGATATCGGAGGATCTCAAGAGTCTCTGAGAAACAGTG 698  
 QY 819 GAGTCTCTCTGAGAGGCGGCTTGAACCTCTCACTTCTCTGAGACCTCTCATTTATTT 878  
 DE 699 GAGTCTCTCTGAGAGGCGGCTTGAACCTCTCACTTCTCTGAGACCTCTCATTTATTT 759  
 QY 879 CAACATCTCCAGGCGAGTTACAGATGTGATTCTGATCAAGCGGCGTGTCTCTCTGGC 938  
 DB 759 CAACATCTCCAGGCGAGTTACAGATGTGATTCTGATCAAGCGGCGTGTCTCTCTGGC 818  
 QY 939 AACGTTTCTACACCTTATACAGACATCCCTGGGCTGCGAGTCTGCGCTATACAGTCT 998  
 DB 819 AACGTTTCTACACCTTATACAGACATCCCTGGGCTGCGAGTCTGCGCTATACAGTCT 878  
 QY 999 TGACATTCGCACTGTTTTTACTTGGGAGATTCAAGGAACAGAAAGTCTCTGATTCACCTG 1058  
 DB 879 TGACATTCGCACTGTTTTTACTTGGGAGATTCAAGGAACAGAAAGTCTCTGATTCACCTG 938  
 QY 1059 GACACCACTTCTGATGAACAGGTTCTTAAGCCAGGCGAGGTTGCTGCTCGCTGCATC 1118  
 DB 939 GACACCACTTCTGATGAACAGGTTCTTAAGCCAGGCGAGGTTGCTGCTCGCTGCATC 998  
 QY 1119 CTCCTTAGAAGATATGAACCTCAATGACTTCCCTGATGATACCGTGAACCTCATCAA 1178  
 DB 999 CTCCTTAGAAGATATGAACCTCAATGACTTCCCTGATGATACCGTGAACCTCATCAA 1058  
 QY 1179 GACGACCGGCTCATGATGAGGAGTCCCTTCCATCTTCAACAGGCGATGGTTCCTGAG 1238







```
QY 1833 CTACCTGACAGACAGACAGACCTTTGGGGGAGTGTCTTCCGTAATATCATCCACAGACAGAA 1892
DB 1779 CAGACGGAGCTCCGGTGGCCACAGAGGGGTCTGGAAAGATCCCAACTGCAAGGCCA 1838
QY 1893 GCGAGTATTCGGGAAGTACCTCAAGAGCCAGACGAGCTGGTCCCGTCAGCCTCTT 1952
DB 1839 GAGAGTATTCGGGAAGTACCTCAAGAGCCAGACGAGCTGGTCCCGTCAGCCTCTT 1898
QY 1953 GGCACTTGCAGTATCTGCTTTCGATTCGAGGGGCTTCTTCTCGGGATCATCAGCTCTA 2012
DB 1899 GGCACTTGCAGTATCTGCTTTCGATTCGAGGGGCTTCTTCTCGGGATCATCAGCTCTA 1958
QY 2013 CTGGGCTGTGATCATGGGCGCAAGACCTGGTGTGTTGTCAGACGGCAAGAGAGAGCT 2072
DB 1959 CTGGGCTGTGATCATGGGCGCAAGACCTGGTGTGTTGTCAGACGGCAAGAGAGCT 2018
QY 2073 CAGCCACTCGCGCGGGCTTCCATGAGCAGCGCTGACCAAGCTCAGCGGCTCTTGGGA 2132
DB 2019 CAGCCACTCGCGCGGGCTTCCATGAGCAGCGCTGACCAAGCTCAGCGGCTCTTGGGA 2078
QY 2133 CAGTCAATCCAAAGACCCAAAGCGGGAGGACATCCCTACAGGACTCATGACAGAGCAA 2192
DB 2079 CAGTCAATCCAAAGACCCAAAGCGGGAGGACATCCCTACAGGACTCATGACAGAGCAA 2138
QY 2193 GCTCGCCACTCCCGGGAACAGCGGCAAGATGCTCATTAAAGACAGCCAGCACCCTGGA 2252
DB 2139 GCTCGCCACTCCCGGGAACAGCGGCAAGATGCTCATTAAAGACAGCCAGCACCCTGGA 2198
QY 2253 CCTGACGGCCCTCCCAACCCAGAGTCAACCCCAAGCCTGCACAGAGCGGAAGCCCA 2312
DB 2199 CCTGACGGCCCTCCCAACCCAGAGTCAACCCCAAGCCTGCACAGAGCGGAAGCCCA 2258
QY 2313 CCGGGCAGCGGGAGTGGGAGGAGAACCAAGACCTCATCATATGCTGCACAAAGACAT 2372
DB 2259 CCGGGCAGCGGGAGTGGGAGGAGAACCAAGACCTCATCATATGCTGCACAAAGACAT 2318
QY 2373 GCGCCCATGGGCTCCCTGTGATGCCAGGACCTGCCCTGCGGGCTTCCCGCCAGCCA 2432
DB 2319 GCGCCCATGGGCTCCCTGTGATGCCAGGACCTGCCCTGCGGGCTTCCCGCCAGCCA 2378
QY 2433 CATCCCGAGCGTGTGCTTCCCTGCCATCAGCGAGCAGGCTACCCATGAGTACGTGGA 2492
DB 2379 CATCCCGAGCGTGTGCTTCCCTGCCATCAGCGAGCAGGCTACCCATGAGTACGTGGA 2438
QY 2493 CCAGCCAAATAGGCGAGTGGCGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 2552
DB 2439 CCAGCCAAATAGGCGAGTGGCGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 2498
QY 2553 GTATAGACATCAAGGAACATCTCAGCAGCAAGAGTCCGCAACCATGGGTGAACCTTCT 2612
DB 2499 GTATAGACATCAAGGAACATCTCAGCAGCAAGAGTCCGCAACCATGGGTGAACCTTCT 2558
QY 2613 GGAGAACCTGGACAGCTGCCCGCAAGTTCACAGCGGAGGAGTCCGTCGGTCCCGC 2672
DB 2559 GGAGAACCTGGACAGCTGCCCGCAAGTTCACAGCGGAGGAGTCCGTCGGTCCCGC 2618
QY 2673 GGGAGCTCCCTCTCTCAGACCTGTATAGCAAGCGCTGCAATGCAACATCTCTCTCT 2732
DB 2619 GGGAGCTCCCTCTCTCAGACCTGTATAGCAAGCGCTGTATAGCAAGCGCTGTCTCTCT 2678
QY 2733 CTAGGGGTTGACTATAGAGGAGCTACCCGCAACTCGCTCGTCCAGAGAGCCAGCAGC 2792
DB 2679 CTAGGGGTTGACTATAGAGGAGCTACCCGCAACTCGCTCGTCCAGAGAGCCAGCAGC 2736
QY 2793 CACCAGCTCTCAAAAGAACAGACTACTCTCTCAATTCCTTCACCTCTCAGAAACCA 2852
DB 2739 CACCAGCTCTCAAAAGAACAGACTACTCTCTCAATTCCTTCACCTCTCAGAAACCA 2798
QY 2853 GAGCTTGGCAGGGAGACAAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2912
DB 2799 GAGCTTGGCAGGGAGACAAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2858
```

```
QY 2913 GCACAGCTCCAGCCCACTCCGACAGCGGTGACTGTCTCGAGGACGCCAGCCTCAACG 2972
DB 2859 GCACAGCTCCAGCCCACTCCGACAGCGGTGACTGTCTCGAGGACGCCAGCCTCAACG 2918
QY 2973 CTAAACTCACTGACAGAGTCGGGGCTGAAGCGTACGCCCTCGCTAAAGCGGACGTACC 3032
DB 2919 CTAAACTCACTGACAGAGTCGGGGCTGAAGCGTACGCCCTCGCTAAAGCGGACGTACC 2978
QY 3033 CCCCAAGCATCTTCTGCTCCCTTTCACATCATCATGAGCCCAATGATCGTGTACATA 3092
DB 2979 CCCCAAGCATCTTCTGCTCCCTTTCACATCATCATGAGCCCAATGATCGTGTACATA 3038
QY 3093 A 3093
DB 3093 A 3093

RESULT 10
AAH17625
ID AAH17625 standard: cDNA; 304; BP.
XX
AC AAH17625;
XX
DI 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:17153.
KW human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI07461.7-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-01:6126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HEP-) HELIX RES INST.
XX
PI Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WTI: 2001-318749/34.
XX
PR Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
PS Claim 8: SEQ ID 17153; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
```



CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13629 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX  
 SQ Sequence 3041 BP; 804 A; 886 C; 714 G; 637 T; 0 other;

Query Match 70.6%; Score 2182.4; DB 22; Length 3041;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 910 CGTATCAACGGCGGTGATGTTGCTCGCAAGG-TTTCTACACGTTATAACAGCATCCCT 969  
 DB 1 CGTATCAACGGCGGAIGATGTTGCTCGCAAGG-TTTCTACACGTTATAACAGCATCCCT 60  
 QY 970 GGGTCTGACATCTGTGCTATGACATCTTGACATTCACAGTGTTTTACTGGGAGATTC 1029  
 DB 61 GGGTCTGACATCTGTGCTATGACATCTTGACATTCACAGTGTTTTACTGGGAGATTC 120  
 QY 1030 AAGGAACAGAGTCTCTGATTCACCTGAGACACCAGTCTCTGATGACAGTTCCTTAAG 1069  
 DB 121 AAGGAACAGAGTCTCTGATTCACCTGAGACACCAGTCTCTGATGACAGTTCCTTAAG 180  
 QY 1090 CCCAGGCGAGTTCCTGTGCTGGCTGATCTCTCTAGAAAGATATGCAAGCTCCAAAGAG 1149  
 DB 181 CCCAGGCGAGTTCCTGTGCTGGCTGATCTCTCTAGAAAGATATGCAAGCTCCAAAGAG 240  
 QY 1150 TTCCCTGATGATACCTCTGATTTTCAAGACAGCAGCCGCTGATGATGAGGAGTCCCT 1209  
 DB 241 TTCCCTGATGATACCTCTGATTTTCAAGACAGCAGCCGCTGATGATGAGGAGTCCCT 300  
 QY 1210 TCCATCTTCAACAGGCGATGTTCTGAGAGAGATGCTGAGATACCGCTTACCAAAAT 1269  
 DB 301 TCCATCTTCAACAGGCGATGTTCTGAGAGAGATGCTGAGATACCGCTTACCAAAAT 360  
 QY 1270 GCAGTGGACAGAGTCTGTGGGCCATATCAATACACACTGTGGTTTCTGGGATCAGAG 1329  
 DB 361 GCAGTGGACAGAGTCTGTGGGCCATATCAATACACACTGTGGTTTCTGGGATCAGAG 420  
 QY 1330 AAGGGAATCATCTGATTTTGGCCAGATATAGAAATATGTTTCTAAATGACAGC 1389  
 DB 421 AAGGGAATCATCTGATTTTGGCCAGATATAGAAATATGTTTCTAAATGACAGC 480  
 QY 1390 CTTTCTGGAGAGATCAGTGTGTTAGTACATCTGAGAAATGCAAGTATGATGAGTCCAA 1449  
 DB 481 CTTTCTGGAGAGATCAGTGTGTTAGTACATCTGAGAAATGCAAGTATGATGAGTCCAA 540  
 QY 1450 GACAAAAGGATCATGGGCAATGACGCTGACAGATCAAGCAGCTCTGTATGTTGGCTTC 1509  
 DB 541 GACAAAAGGATCATGGGCAATGACGCTGACAGATCAAGCAGCTCTGTATGTTGGCTTC 600  
 QY 1510 TCTACCTGTGTATAAGGTTCCTCTGGCGGTGTGAACACATGCGAGTGTAAATAA 1569  
 DB 601 TCTACCTGTGTATAAGGTTCCTCTGGCGGTGTGAACACATGCGAGTGTAAATAA 660  
 QY 1570 ACCTGTATGCTCCAGAGACCCATATGTTGGATGATTAAGGAAGTGTGCTCGAGC 1629  
 DB 661 ACCTGTATGCTCCAGAGACCCATATGTTGGATGATTAAGGAAGTGTGCTCGAGC 720  
 QY 1630 CATTTATCAACCAACAGCAGACTGACITTTGACGAGACATAGCGTGGCAATACAGAT 1689  
 DB 721 CATTTATCAACCAACAGCAGACTGACITTTGACGAGACATAGCGTGGCAATACAGAT 750  
 QY 1690 GGCTCGGGGACTGTCAAAATTCCTTTGTGGCACTGAATGGGATTCAGATTCCTCTC 1749  
 DB 781 GGCTCGGGGACTGTCAAAATTCCTTTGTGGCACTGAATGGGATTCAGATTCCTCTC 840  
 QY 1750 CCCAGCAACACCAATCAGATTCGACGGCTCAAGAGGGGATGATCTAGGGAGGATG 1809  
 DB 841 CCCAGCAACACCAATCAGATTCGACGGCTCAAGAGGGGATGATCTAGGGAGGATG 900

QY 1810 CTGGACTGGAAGCAATCGCTTGACTCACCIGACAGACAGACCCCTTTGGGGGAGTGTCT 1869  
 DB 901 CTGGACTGGAAGCAATCGCTTGACTCACCIGACAGACAGACCCCTTTGGGGGAGTGTCT 960  
 QY 1870 TCCCATAAATCACCACAAAGAGGAGTGTATCGGGAAGTACTCTCAAGAGCCACAGAC 1929  
 DB 961 TCCCATAAATCACCACAAAGAGGAGTGTATCGGGAAGTACTCTCAAGAGCCACAGAC 1020  
 QY 1930 CAGCTGGTTCCCGTCAACCTCTTTGGCCATTGCGAGTCATCTCGCTTTCCTCAIGGGGCC 1989  
 DB 1021 CAGCTGGTTCCCGTCAACCTCTTTGGCCATTGCGAGTCATCTCGCTTTCCTCAIGGGGCC 1080  
 QY 1990 GTCTTCTCGGCGATCACCGTCTACCTGCTGATCATCGGCGCAAGACAGTGGCTGTG 2049  
 DB 1081 GTCTTCTCGGCGATCACCGTCTACCTGCTGATCATCGGCGCAAGACAGTGGCTGTG 1140  
 QY 2050 GTGACGCGAAGAGAGAGGAGTCAACCCACTCGCGCGGGCTCCATGATGACAGCTCAC 2109  
 DB 1141 GTGACGCGAAGAGAGAGGAGTCAACCCACTCGCGCGGGCTCCATGATGACAGCTCAC 1200  
 QY 2110 AAGCTCAGGGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAGGCGCATCTC 2169  
 DB 1201 AAGCTCAGGGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAGGCGCATCTC 1260  
 QY 2170 AGCCCACTCATGACAAACGCAAGCTTCGCCACTTCGCCGCAACACGCGCAAGATGCTCAT 2229  
 DB 1261 AGCCCACTCATGACAAACGCAAGCTTCGCCACTTCGCCGCAACACGCGCAAGATGCTCAT 1320  
 QY 2230 AAAGCAGACAGCACCATCTGACCTGAGGGGCTCTCCACCCAGAGTCAACCCCAAG 2289  
 DB 1321 AAAGCAGACAGCACCATCTGACCTGAGGGGCTCTCCACCCAGAGTCAACCCCAAG 1380  
 QY 2290 GTGACGACAGAGGGAAGCGCGCGCGAGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAG 2349  
 DB 1381 GTGACGACAGAGGGAAGCGCGCGCGAGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
 QY 2350 ATCAATCCCTGCACAAAGGACATGCCCCCATGCGGTCCCTCTGTATTCACACGAGCCTG 2409  
 DB 1441 ATCAATCCCTGCACAAAGGACATGCCCCCATGCGGTCCCTCTGTATTCACACGAGCCTG 1500  
 QY 2410 CCCTCTGGGCGCTCCCGCAGGACATCCCGAGCTGTGGTCTCTGCTCCATCAGCAGAG 2469  
 DB 1501 CCCTCTGGGCGCTCCCGCAGGACATCCCGAGCTGTGGTCTCTGCTCCATCAGCAGAG 1560  
 QY 2470 GGCTACACAGTACGATGAGTGGACAGCCCAAAATGAGCGAGTGGCCAGATGGCGTG 2529  
 DB 1561 GGCTACACAGTACGATGAGTGGACAGCCCAAAATGAGCGAGTGGCCAGATGGCGTG 1620  
 QY 2530 GAGGACAGGCGGCCACACTGGAGTATAAGACCATCAAGGAACATCTACGACAGAGT 2589  
 DB 1621 GAGGACAGGCGGCCACACTGGAGTATAAGACCATCAAGGAACATCTACGACAGAGT 1680  
 QY 2590 CCAACCATGGGGTGAACCTTGTGGAGACCTGGAGAGCTGCCCGCCCAAGATTCCACAG 2649  
 DB 1681 CCAACCATGGGGTGAACCTTGTGGAGAACCTTGTGGAGAACCTGGAGAGCTGCCCGCCCAAGATTCCACAG 1740  
 QY 2650 CGGAGGCGCTCCCTGGGTCCCGCGGAGCTCCCTGTCTCAGACGGTCTAGCAGAGCG 2709  
 DB 1741 CGGAGGCGCTCCCTGGGTCCCGCGGAGCTCCCTGTCTCAGACGGTCTAGCAGAGCG 1800  
 QY 2710 CTGGAATGCACACTCTCTTCTACGGGGTGTACTATAAGAGGAGCTACCCACAGAC 2769  
 DB 1801 CTGGAATGCACACTCTCTTCTACGGGGTGTACTATAAGAGGAGCTACCCACAGAC 1860  
 QY 2770 TGCTCAGAGAGCCACAGGCGCCACTCTCAAAAGAAACAACTAACTCTCTCCAT 2829  
 DB 1861 TGCTCAGAGAGCCACAGGCGCCACTCTCAAAAGAAACAACTAACTCTCTCCAT 1920  
 QY 2830 TCCCTCTCACTCTCCAGAAACAGAGCTTTGCGAGGGAGACAAACCGCGCGCGCGCGCG 2889  
 DB 1921 TCCCTCTCACTCTCCAGAAACAGAGCTTTGCGAGGGAGACAAACCGCGCGCGCGCGCG 1980  
 QY 2890 CAGAGGGTGGACTCCATCCAGGTGCACAGCTCCACGCCATCTGGCAGGCCCTGACTGTC 2949

Db 1981 CAGAGGTGACCTCCATCCAGTGGACAGCTCCAGCCATCTCCGACGACCCGCTGCTGTC 2040  
 Qy 2950 TCGAGGAGCCGAGCCCTCAAGGCTCAAGCTCACTGACAGGTGGGGTGAAGCGTAGG 3009  
 Db 2041 TCGAGGAGCCGAGCCCTCAAGGCTCAAGCTCACTGACAGGTGGGGTGAAGCGTAGG 2100  
 Qy 3010 CCTCGGTAAAGCGGAGCTACCCGACCAACCATCTTTGCTCCCTTTCCAGTCCATG 3069  
 Db 2101 CCTCGGTAAAGCGGAGCTACCCGACCAACCATCTTTGCTCCCTTTCCAGTCCATG 2160  
 Qy 3070 AAGCCCAATGATGGTACATATA 3093  
 Db 2161 AAGCCCAATGATGGTACATATA 2184

RESULT 11  
 ID AAA93630 standard: cDNA; 1840 BP.  
 XX  
 AC AAA93630;  
 XX  
 DT 16-JAN-2001 (first entry)  
 DE Human scaphorin protein-like splice variant SECX pCR2.1-2864933 cDNA.  
 XX  
 KW SECX protein; human; secreted; membrane-associated; cancer;  
 KW proliferation regulator; differentiation regulator; non-malignant tumour;  
 KW immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
 KW infection; inflammatory disorder; arthritis; haematopoietic disorder;  
 KW skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
 KW neurological disease; Alzheimer's disease; trauma; wounding;  
 KW spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
 KW anti-HIV; anti-inflammatory; antiarthritic; antiarteriosclerotic;  
 KW neuroprotective; vulnerary; antiallergic; antimicrobial; cardiatic;  
 KW dermatological; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200053742-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PP 09-MAR-2000; 2000WO-US06280.  
 XX  
 PR 09-MAR-1999; 99US-0123667.  
 PR 08-MAR-2000; 2000US-0123667.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA;  
 XX  
 WP: 2000-5943:2/56.  
 DR P-PSDB: AAB23043.  
 XX  
 PT Novel human membrane associated or secreted polypeptides and  
 PT polynucleotides useful for diagnosis, prevention and treatment of  
 PT pathological states such as cancer, immune, cardiovascular and  
 PT neurological disorders.  
 XX  
 PS Claim 3; Fig 15; 151pp; English.  
 XX  
 CC Sequences AAA93616-A93631 and AAA93673-A93676 represent nucleic acids  
 CC which encode human SECX proteins (AAB23029-B3045). The SECX proteins  
 CC of the invention are either secreted or membrane-associated proteins  
 CC and act as regulator of cellular proliferation and differentiation. SECX  
 CC proteins or nucleotides are useful for diagnosing the presence of, or  
 CC predisposition to, a disease associated with altered levels of SECX  
 CC proteins and nucleotides. The SECX proteins are also useful to screen  
 CC compounds that modulate SECX activity or expression. The interaction of  
 CC a SECX protein with other cellular proteins may be useful to modulate  
 CC the activity of a partner protein, cellular proliferation, cellular  
 CC differentiation and cell survival. SECX nucleotides are useful for the

recombinant expression of SECX protein, and may be used to detect SECX mRNA  
 or genetic lesions in the SECX gene. They may also be used to modulate  
 SECX expression (e.g., using antisense oligonucleotides). SECX nucleic  
 acid sequences are also useful for identifying a cell or tissue type in  
 a biological sample, and in forensic biology. SECX primers or probes are  
 useful for detecting the presence of SECX nucleotides and for screening  
 tissue cultures for contamination. Diseases that may be treated or  
 prevented using SECX proteins or nucleotides include cancer (e.g.,  
 colorectal carcinoma, prostate cancer), benign tumours, immune disorders  
 (including autoimmune diseases, transplant rejection, allergies, AIDS),  
 infections, inflammatory disorders, arthritis, haematopoietic disorders,  
 skin disorders, cardiovascular disorders, atherosclerosis, restenosis,  
 neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,  
 surgical or traumatic wounds, spinal cord injury), and skeletal  
 disorders.  
 XX  
 SQ Sequence 1890 BP; 535 A; 426 C; 463 G; 466 T; 0 other;  
 Query Match 60.8%; Score 1881.6; DB 21; Length 1890;  
 Best Local Similarity 95.8%; Pred. No. 0;  
 Matches 1884; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 Qy 49 GGGGCTGGTTTCCGAGAGATTCTCAGCCAACTAGTATTTCGATGGCACTATACAAA 108  
 Db 1 GGA:CCGG:TTCCGAGAAGATTCAGCCAACTAG:ATTTCCGATGGCACTATACAAA 60  
 Qy 109 CAGTATCCGGTGTGTGGGCCCAAGCCAGGAGGGAACCCACACAGAGGCGACAGGCTG 168  
 Db 61 CAGTATCCGGTGTGTGGGCCCAAGCCAGGAGGGAACCCACACAGAGGCGACAGGCTG 120  
 Qy 169 GACATCCAGATGATTATGATCATGAACGGAACTCTTACATTCGCTGTAGGACCATATT 228  
 Db 121 GACATCCAGATGATTATGATCATGAACGGAACTCTTACATTCGCTGTAGGACCATATT 180  
 Qy 229 TATACGTGTGATATAGACACATCACACGGGAAGAAATTTATTGTACAAAAGACTGACA 288  
 Db 131 TATACGTGTGATATAGACACATCACACGGGAAGAAATTTATTGTACAAAAGACTGACA 240  
 Qy 289 TGGAAATCTAGACAGGGCGATGTAGACACATGCAAGATGAAGGGAACATAGGATGAG 348  
 Db 241 TGGAAATCTAGACAGGGCGATGTAGACACATGCAAGATGAAGGGAACATAGGATGAG 300  
 Qy 349 TGGCAAACTTTTAAAGTTCTTCTAAAGAAAACGATGATGCATTTGTTGCTGTGGA 408  
 Db 301 TGGCAAACTTTTAAAGTTCTTCTAAAGAAAACGATGATGCATTTGTTGCTGTGGA 360  
 Qy 409 ACTAATGCCCTTCAACCCCTTCCTGCAGAACTATAGATGGATGATCATTTGGACCATTCGG 468  
 Db 361 ACTAATGCCCTTCAACCCCTTCCTGCAGAACTATAGATGGATGATCATTTGGACCATTCGG 420  
 Qy 469 GATGAATTCAGCGGAATGGCCAGATGCCCATATATGATGCCAAACATGCCACGTTGCAC 528  
 Db 421 GATGAATTCAGCGGAATGGCCAGATGCCCATATATGATGCCAAACATGCCACGTTGCAC 480  
 Qy 529 TTTCAGATGGAAAACATATCTCAGCCACAGTGCATCTGCTTCCCTTCATTCAGCGAGTC 588  
 Db 481 TTTCAGATGGAAAACATATCTCAGCCACAGTGCATCTGCTTCCCTTCATTCAGCGAGTC 540  
 Qy 589 ATTACCGCGATCTTGGAGAAAGCCCTACCCCTCGGAGCCGCTCAAGCAGCATTCAAAATGG 648  
 Db 541 ATTACCGCGATCTTGGAGAAAGCCCTACCCCTCGGAGCCGCTCAAGCAGCATTCAAAATGG 600  
 Qy 649 TTGAAGAACCATACTTTTTCAGGCGGTGGATTTACGGAGATATATATCTACTTCTTCTTC 708  
 Db 601 TTGAAGAACCATACTTTTTCAGGCGGTGGATTTACGGAGATATATATCTACTTCTTCTTC 660  
 Qy 709 AGGGAATAGCAGTGGACTATACACCATAGGAAGTAGTTTTTCCCAAGAGTGGCTCAG 768  
 Db 661 AGGGAATAGCAGTGGAGTATACACCATAGGAAGTAGTTTTTCCCAAGAGTGGCTCAG 720  
 Qy 769 GTTTGTAGAAGATGATGGGAGGATCTCAAGAGTCTCTGGAGAAACAGTGGAGCTCTTC 828  
 Db 721 GTTTGTAGAAGATGATGGGAGGATCTCAAGAGTCTCTGGAGAAACAGTGGAGCTCTTC 780





PR 08-JUL-1999; 95JP-0194486.  
PR 11-JAN-2000; 2000JP-0118774.  
PR 02-MAY-2000; 2000JP-0183765.  
XX  
XX (HELI-) HELIX RES INST.  
XX Ota T, Nishikawa T, Isoyui T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
XX WP1: 2001-524255/58.  
DR P-PSDB: NAM93444.  
XX  
XX 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
XX  
XX Claim 8: SEQ ID NO 3087; 1380pp + sequence listing; English.  
XX  
XX The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesizing the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a full length  
CC human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX  
XX Sequence 2227 BP: 567 A; 760 C; 566 G; 460 T; 0 other:  
  
Query Match 56.3%; Score 1742.4; DB 22; Length 2227;  
Best Local Similarity 99.9%; Pred. No. C;  
Matches 1743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1350 TTTGGCCAGATAGGAATAGTGGTITCTTAATGACACAGCCCTTTCCGAGGAGATGAG 1403  
Db 1 TTTGGCCAGATAGGAATAGTGGTITCTTAATGACACAGCCCTTTCCGAGGAGATGAG 60  
  
QY 141C TGTTTACACTCTGAAAATCAGCTATGATGAGTGAAGACAAAGGATCATGGGCAT 1469  
Db 61 TGTTTACACTCTGAAAATCAGCTATGATGAGTGAAGACAAAGGATCATGGGCAT 120  
  
QY 1470 GCAGCTGGACAGCAGACAGCTCTCTGATGTTCGGTTCTCTACCTGTGTGATAAGGCT 1529  
Db 121 GCAGCTGGACAGCAGACAGCA KTCCTCTGATGTTCGGTTCTCTACCTGTGTGATAAGGCT 130  
  
QY 1530 TCCCTTGGCCGGTGTGAAACACATGAGTGGGAAGTGTAAATAAGCTGTATGCCCTCCAGAGA 1589  
Db 181 TCCCTTGGCCGGTGTGAAACACATGAGTGGGAAGTGTAAATAAGCTGTATGCCCTCCAGAGA 240  
  
QY 1590 CCATATCTGGATGGATAAGGAAGGTGGTGGCTGAGCCATTTATCAGCCACACAGAG 1649  
Db 241 CCATATCTGGATGGATAAGGAAGGTGGTGGCTGAGCCATTTATCAGCCACACAGAG 300  
  
QY 1650 ACTGACTTTTGACGAGGACATAGACGCTGGCAATACAGATGGCTTGGGGGACTGTCAAA 1709  
Db 301 ACTGACTTTTGACGAGGACATAGACGCTGGCAATACAGATGGCTTGGGGGACTGTCAAA 360  
  
QY 1710 TTCCCTTGGGACTGAATGGGCATTCAGTTCCTCTTGGCCAGCACAACCATCAGA 1769  
Db 361 TTCCCTTGGGACTGAATGGGCATTCAGTTCCTCTTGGCCAGCACAACCATCAGA 420  
  
QY 1770 TTCAGCGCTCAAGAGGGGTATGAGTCTAGGCGAGGAATGTGGAGCTGGAAGCTCTGCT 1829  
Db 421 TTCAGCGCTCAAGAGGGGTATGAGTCTAGGCGAGGAATGTGGAGCTGGAAGCTCTGCT 480  
  
QY 1830 TGACTCACCTGACAGACAGACCCCTTTGGGGGAGTGTCTTCCCAATATCAACCAAGACAA 1889  
Db 481 TGACTCACCTGACAGACAGACCCCTTTGGGGGAGTGTCTTCCCAATATCAACCAAGACAA 540  
  
QY 1890 GAAGGAGTGTATTCGGGAAAGTTACTTCAAGGCCACAGCACCTGGTTCGGGTCAACCT 1949

Db 541 GAAGGAGTGTATTCGGGAAAGTTACTTCAAGGCCACAGCAGCTGGTTCGGTCAACCT 600  
QY 1950 CTTGGCCATTGACATCTGGCTTTCTGTCATGGGGCCGCTCTCTCGGGGATCACCT 2009  
Db 601 CTTGGCCATTGACATCTGGCTTTCTGTCATGGGGCCGCTCTCTCGGGGATCACCT 660  
QY 2010 CTACTCGCTCTGTGATCATCGGCGCAAGACGTGTGTGTGTCAGCGCAAGGAGAA 2069  
Db 661 CTACTCGCTCTGTGATCATCGGCGCAAGACGTGTGTGTGTCAGCGCAAGGAGAA 720  
QY 2070 GCTACCCACTCGCGCGGGCTCCATGAGCAGCTCACCAAGCTCAGCGGCTCTTTGG 2129  
Db 721 GCTACCCACTCGCGCGGGCTCCATGAGCAGCTCACCAAGCTCAGCGGCTCTTTGG 780  
QY 2130 GGACACTCAATCCMAAGACCCAAAGCCGAGGCTCATCTCAGCCACTCATGCAACAGG 2189  
Db 781 GGACACTCAATCCMAAGACCCAAAGCCGAGGCTCATCTCAGCCACTCATGCAACAGG 840  
QY 2190 CAACTCGCCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGCAGCAGCAGCT 2249  
Db 841 CAACTCGCCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGCAGCAGCAGCT 900  
QY 2250 GGACTGAGCGGCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGCAGCAGCAG 2309  
Db 901 GGACTGAGCGGCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGCAGCAGCAG 960  
QY 2310 CAGCCGCGGAGCGCGGAGTGGAGAGAACAGACCTCATCAATGCTGCAACAGAA 2369  
Db 961 CAGCCGCGGAGCGCGGAGTGGAGAGAACAGACCTCATCAATGCTGCAACAGAA 1020  
QY 2370 CATGCCCGCATGGGCTCCCGTCTGATTCGCCAGGACTGCGGCTGGCGGCTCCCGCAG 2429  
Db 1021 CATGCCCGCATGGGCTCCCGTCTGATTCGCCAGGACTGCGGCTGGCGGCTCCCGCAG 1080  
QY 2430 CCACATCCCGAGGCTGTCTCTGCTCCCATCAGCAGCAGGCTACACATGAGTACCT 2489  
Db 1081 CCACATCCCGAGGCTGTCTCTGCTCCCATCAGCAGCAGGCTACACATGAGTACCT 1140  
QY 2490 GGACGAGCCCAATGAGCGAGTGGCGGAGTGGCGGCTGGAGGAGCAGCGCGCACCT 2549  
Db 1141 GGACGAGCCCAATGAGCGAGTGGCGGAGTGGCGGCTGGAGGAGCAGCGCGCACCT 1200  
QY 2550 GGAGTATAAGACCATCAAGGAACTCTCAGCAGCAAGTGTCCCAACCATGGGCTGAACCT 2609  
Db 1201 GGAGTATAAGACCATCAAGGAACTCTCAGCAGCAAGTGTCCCAACCATGGGCTGAACCT 1260  
QY 2610 TGTGGAGAGCTGGACAGCTGCGCGGCAAGTTCACAGCGGAGGCTGCTGCTGGTCC 2669  
Db 1261 TGTGGAGAGCTGGACAGCTGCGCGGCAAGTTCACAGCGGAGGCTGCTGCTGGTCC 1320  
QY 2670 CCGGAGAGCTGCTGCTGCTCAGAGCTGGTCTAAGCAAGCGGCTGGAATGCACCTCCCTC 2729  
Db 1321 CCGGAGAGCTGCTGCTGCTCAGAGCTGGTCTAAGCAAGCGGCTGGAATGCACCTCCCTC 1380  
QY 2730 TTCTTACGGGCTGTGATATTAAGAGAGCTACCCGCAAGCTCGCTCAGGAGGAGCAGCA 2789  
Db 1381 TTCTTACGGGCTGTGATATTAAGAGAGCTACCCGCAAGCTCGCTCAGGAGGAGCAGCA 1440  
QY 2790 GGCCACGACTCTCAAGAGAGAGAGCTACTCTCCATTCCTCAGCTCTCCAGAGAA 2849  
Db 1441 GGCCACGACTCTCTCAAGAGAGAGAGCTACTCTCCATTCCTCAGCTCTCCAGAGAA 1500  
QY 2850 CCAGAGCTTTTGGCAGGAGGAGCAACCCGCGCGCGCGCGGAGGCTGGAGCTCCATCCA 2909  
Db 1501 CCAGAGCTTTTGGCAGGAGGAGAGAGCCGCGCGCGCGCGCGGAGGCTGGAGCTCCATCCA 1560  
QY 2910 GGTGACAGCTTCCAGCCATCTCGCGAGGCGGTGAGTGTCTCGAGGAGCAGCCAGCCTCAA 2969  
Db 1561 GGTGACAGCTTCCAGCCATCTCGCGAGGCGGTGAGTGTCTCGAGGAGCAGCCAGCCTCAA 1620  
QY 2970 CGCTTACACTCACTGACAGAGTGGGGCTGAGCGTACGCTCGCTCGCTTAAGCGGAGCT 3029

Db 1621 CGCTACAACTCACTGACANGTGGGGGTGAAGGACGACGCTGCTAAAGCGGAGCT 1620  
 QY 3030 ACCCCCAACCACTCTTGGCTCCCTTTCCATCATCAAGCCCAATGATGCTGAC 3089  
 Db 1661 ACCCCCAACCACTCTTGGCTCCCTTTCCATCATCAAGCCCAATGATGCTGAC 1749  
 QY 3090 ATAA 3093  
 Db 1741 ATAA 1744  
 RESULT 14  
 AAH17861  
 ID AAH17861 standard; cDNA: 2306 BP.  
 AC AAH17861;  
 XX 26-JUN-2001 (first entry)  
 DE Human cDNA sequence SEQ ID NO:17567.  
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KW Homo sapiens.  
 OS EP1074517-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX (HELI-X) HELIX RES INST.  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI: 2001-318749/34.  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PI full-length cDNAs -  
 XX Claim 8; SEQ ID 17567; 2537pp + CD ROM; English.  
 XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification, where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification; where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily under specialized methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAB35893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 2306 BP: 579 A: 739 C: 577 G: 411 T: 0 other:  
 SO Query Match: 53.9%; Score 1666.4; DB 22; Length 2306;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1667; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1426 AAATGACGCTATGATGGAGTCGAAGACAAAGGATCATGGGCATGAGCTGACAGACGA 1485  
 Db 166 AGATGCAGCTATGATGAGTCGAAGACAAAGGATCATGGGCATGAGCTGACAGACGA 225  
 QY 1485 AGCAGCTCTCTGATGTTGGCTTCTACCTGTGTGAIAAAGGTCCTCTGGCCGGTGT 1545  
 Db 226 AGCAGCTCTCTGATGTTGGCTTCTACCTGTGTGAIAAAGGTCCTCTGGCCGGTGT 285  
 QY 1545 GAAGGACATGGGAAGTGTAAAGAAAGCTGTATGCTCCAGAGACCCATATTTGGATGG 1605  
 Db 286 GAAGGACATGGGAAGTGTAAAGAAAGCTGTATGCTCCAGAGACCCATATTTGGATGG 345  
 QY 1606 ATRAGGAAGTGTGGCTGAGGCTATATACCCAGACAGACGACTGCTTTGAGCAG 1665  
 Db 346 ATAAGGAAGTGTGGCTGAGGCTATATACCCAGACAGACGACTGCTTTGAGCAG 405  
 QY 1666 GACATGAGGCTGSCAATACAGATGGTCTGGGGAGCTGTACAAATTCCTTTGTGGCACTG 1725  
 Db 406 GACATGAGGCTGSCAATACAGATGGTCTGGGGAGCTGTACAAATTCCTTTGTGGCACTG 465  
 QY 1725 AATGGGATTCAGTTCCCTCTTCCAGACAGACACACATCAGATTCGACGGGCTCAAGAG 1785  
 Db 466 AATGGGATTCAGTTCCCTCTTCCAGACAGACACACATCAGATTCGACGGGCTCAAGAG 525  
 QY 1786 CGGTATGAGTCTAGGGGAGGAATGCTGGACTGGAAGCATCTGCTTGACTCACCTGACAGC 1845  
 Db 526 CGGTATGAGTCTAGGGGAGGAATGCTGGACTGGAAGCATCTGCTTGACTCACCTGACAGC 585  
 QY 1845 ACAGACCCCTTTGGGGGAGTGTCTCCCATTAATACCAAGACAGAGGAGTGTATCGG 1905  
 Db 586 ACAGACCCCTTTGGGGGAGTGTCTCCCATTAATACCAAGACAGAGGAGTGTATCGG 645  
 QY 1906 GAAAGTTACCTCAAGGCCAGCAGCAGCTGGTTCCCGTCACCTCTTTGGGCATTTGAGCTC 1965  
 Db 646 GAAAGTTACCTCAAGGCCAGCAGCAGCTGGTTCCCGTCACCTCTTTGGGCATTTGAGCTC 705  
 QY 1966 ATCTGGCTTTGCTATGSGGGGCGCTTCTCGGGCATCACCTCTACTGCTCTGTGAT 2025  
 Db 706 ATCTGGCTTTGCTATGSGGGGCGCTTCTCGGGCATCACCTCTACTGCTCTGTGAT 765  
 QY 2026 CATCGGCGCAAGACGTTGGTGTGGTGCAGCGGCAAGGAGAGGAGCTCAGCCACTCGGC 2085  
 Db 766 CATCGGCGCAAGACGTTGGTGTGGTGCAGCGGCAAGGAGAGGAGCTCAGCCACTCGGC 825  
 QY 2086 CUGGGCTCCATGACGACGCTCACCAGCTCAGGGGCTCTTTGGGGACACTCAATCCAA 2145  
 Db 826 CUGGGCTCCATGACGACGCTCACCAGCTCAGGGGCTCTTTGGGGACACTCAATCCAA 885  
 QY 2146 GACCCAAAGCGGAGGCGCATCTCTACGCCACTATGCAACAAGGAGGAGTGGCCACTGCC 2205  
 Db 886 GACCCAAAGCGGAGGCGCATCTCTACGCCACTATGCAACAAGGAGGAGTGGCCACTGCC 945  
 QY 2206 GGCACACGGGCAAGATGCTCATTAAGAGCAGACACAGCCACCTGGACCTGAGGGGCGCTC 2265  
 Db 946 GGCACACGGGCAAGATGCTCATTAAGAGCAGACACAGCCACCTGGACCTGAGGGGCGCTC 1005  
 QY 2266 CCCACCCGAGAGTCAACCCCAACGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2325  
 Db 1006 CCCACCCGAGAGTCAACCCCAACGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1065  
 QY 2326 GATGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2385  
 Db 1066 GATGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1125  
 QY 2386 TCCCTGTGATTCGACGAGCTGCGGCTCGGCGCTCCGCGCAGGACACATCCCGCAGCGTG 2445

1126	TTCCGCTGTGATCTCCACGGGACCTGCCCTGGGGGUGCTCCGCCAGGCCACATCCGAGGGTG	1181
2446	GTGGTCTCTGCCCATCATCAGCAGCAGAGGCTACCAAGCATGAGTACGTGGACAGCCCAAAATG	2505
1186	TTTGGTCTCTGCCCATCATCAGCAGCAGGAGCTACCAAGCATGAGTACGTGGACAGCCCAAAATG	1245
2506	AGCAGAGTGGCCCAATGCGGCTGGAGAGCAGGCGGCACACTGGAGTATAGACCATC	2565
1246	ACCGAGGTGGCCCAATGCGGCTGGAGAGCAGGCGGCACACTGGAGTATAGACCATC	1305
2566	AAGGAACATCTCAGCAGCAASAGTCCCAACCATGGGGTCAACCTTGTCAGAGCACTGGAC	2625
1306	AAGGAACATCTCAGCAGCAASAGTCCCAACCATGGGGTCAACCTTGTCAGAGCACTGGAC	1365
2626	AGCTGGCCCCAAAGTTCCACAGCGGAGGCGCTCCCTGGGTGCCCGGGAGGCTTCCTG	2685
1366	AGCTGGCCCCAAAGTTCCACAGCGGAGGCGCTCCCTGGGTGCCCGGGAGGCTTCCTG	1425
2686	TCTCAGACCGGTCCTAAGCAAGCGGCTGGAAATGCCACACTGCTCTCTACGGGTTGAC	2745
1426	TCTCAGACCGGTCCTAAGCAAGCGGCTGGAAATGCCACACTGCTCTCTACGGGTTGAC	1485
2746	TATAGAGAGAGCTACGCCACCAACTGCTCAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCA	2805
1486	TATAGAGAGAGCTACGCCACCAACTGCTCAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCA	1545
2806	AGAACACACTAATCTCTCAATCTCTCACTCTCTCAGAAACAGAGCTTTTGGCAGG	2865
1546	AGAACACACTAATCTCTCAATCTCTCACTCTCTCAGAAACAGAGCTTTTGGCAGG	1605
2866	GGAGACAACCGCGCGCGCGCGAGAGGTTGAGCTCATCAGTGCACAGCTTCACAGCTCCAG	2925
1606	GGAGACAACCGCGCGCGCGCGAGAGGTTGAGCTCATCAGTGCACAGCTTCACAGCTCCAG	1665
2926	CCATCTGCCAGCGGTGACTGTCTAGGCGAGCGCCACCTCAACGCGCTACAGCTCAGT	2985
1666	CCATCTGCCAGCGGTGACTGTCTAGGCGAGCGCCACCTCAACGCGCTACAGCTCAGT	1725
2986	ACAAGCTGGGGCTGAAGCGTACGCGCTGGCTAAAGCGGAGGTACCGCCCAAAACCATCC	3045
1726	ACAAGCTGGGGCTGAAGCGTACGCGCTGGCTAAAGCGGAGGTACCGCCCAAAACCATCC	1785
3046	TTTGTCGGCTTCCACATCATGAGCGCATGATGCTGGGTGACATAA	3093
1786	TTTGTCGGCTTCCACATCATGAGCGCATGATGCTGGGTGACATAA	1833
RESULT 15		
AAS68253		
ID	AAS68253 standard: cDNA: 2592 BP.	
AC	AAS68253;	
CC		
TT	13-FEB-2002 (first entry)	
XX		
XX	cDNA encoding novel human diagnostic protein: #4057.	
XX		
KW	Human: chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement; medical imaging; diagnostic; genotoxic; genotoxic; ss.	
XX	Homo sapiens.	
XX		
PN	W0200175067-A2.	
XX		
PD	11-OCT-2001.	
XX		
PF	30-MAR-2001; 2001WO-US08631.	
XX		
PR	31-MAR-2000; 2000US-0540217.	
XX		
XX	23-AUG-2000; 2000US-0649167.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		

Drmanac RI, Liu C, Tang YT;  
 WPI: 2001-639362/73.  
 P-PSDB; ABG04066.  
 New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity -  
 Claim 1: SEQ ID No 4057; 103pp; English.  
 The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 and gene mapping, and in recombinant production of (II). The  
 polynucleotides are also used in diagnostics as expressed sequence tags  
 for identifying expressed genes. (I) is useful in gene therapy techniques  
 to restore normal activity of (II) or to treat disease states involving  
 (II). (II) is useful for generating antibodies against it, detecting or  
 quantitating a polypeptide in tissue, as molecular weight markers and as  
 a food supplement. (II) and its binding partners are useful in medical  
 imaging of sites expressing (II). (I) and (II) are useful for treating  
 disorders involving aberrant protein expression or biological activity.  
 The polypeptide and polynucleotide sequences have applications in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits to assess biodiversity  
 and to produce other types of data and products dependent on DNA and  
 amino acid sequences. AAS64197-AAS94564 represent novel human  
 diagnostic coding sequences of the invention.  
 Note: The sequence data for this patent did not appear in the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
 Sequence 2592 BP; 655 A; 754 C; 659 G; 524 T; 0 other;  
 Query Match. 51.5%; Score 1592; DB 23; Length 2592;  
 Best Local Similarity 84.3%; Pred. No. 0;  
 Matches 1809; Conservative 0; Mismatches 325; Indels 12; Gaps 17;  
 QY 960 CAGCATCCCGGGGTCTGAGTCTGTGGCTATGACATGCTTGACATGCCAGTGTITTTAC 1019  
 DB 447 CAGATCCCTGGTGTGCTGCTATGACATGCTTGACATGCCAGTGTITTTAC 506  
 QY 1020 TGGGAGATTCAGGACACAGAGTCTCTGATCTCCCTGGACACACGATTCCTGA:GACG 1079  
 DB 507 TGGGAGATTCAGGACACAGAGTCTCTGATCTCCCTGGACACACGATTCCTGA:GACG 565  
 QY 1080 AGTTCCTTAAGCCAGGCGAGGTCTGCTGTGGTCTCATCTCTCTTACAAGATATGCAAC 1139  
 DB 567 AGTTCCTTAAGCCAGGCGAGGTCTGCTGTGGTCTCATCTCTCTTACAAGATATGCAAC 626  
 QY 1140 CTCCAATGAGTTCCTCATGATACCTTGAACCTTCATCAAGACGACCCGGCTCATGGATGA 1199  
 DB 627 CTCCAATGAGTTCCTCATGATACCTTCATCAAGACGACCCGGCTCATGGATGA 686  
 QY 1200 GGCAGTGCCTTCATCTCTCAACAGGCCATCGTTCCTCGAGACACATGTCAGATACCGCCT 1259  
 DB 687 GGCAGTGCCTTCATCTCTCAACAGGCCATCGTTCCTCGAGACACATGTCAGATACCGCCT 746  
 QY 1260 TACCAGAAATTCAGTGGACACAGCTGCTGGGCCATATCAGAAATCAGACTGTGGTITTTTCT 1319  
 DB 747 TACCAGAAATTCAGTGGACACAGCTGCTGGGCCATATCAGAAATCAGACTGTGGTITTTCT 806  
 QY 1320 GGGATCAGAGGAAATCATCTTTGAAGTITTTGGCCAGAAATAGGAATAGTGGTITTTCT 1379  
 DB 807 GGGATCAGAGGAAATCATCTTTGAAGTITTTGGCCAGAAATAGGAATAGTGGTITTTCT 866  
 QY 1380 AATGACAGCCTTTTCTCGAGGAGATGAGTGTTTACAACCTCTGAAAAATGCAGCTATGA 1439  
 DB 867 AATGACAGCCTTTTCTCGAGGAGATGAGTGTTTACAACCTCTGAAAAATGCAGCTATGA 926  
 QY 1440 TGGAGTCCAGACAAAAGGATCATCTGGGCATGCGAGTGCAGACAGCAAGCAGCTCTCTGTGA 1499



Db 927 CGAAACACCTGTGAGGTTACCAATTATCTTAACCCCTGGTCAAGCTAGTTCTCCATCAC 986  
QY 1500 TGTGGGTTCTTACCTGTGTGATAAAGGTTCCGTTGSCGGTGTGAGGACATGGCAA 1553  
Db 987 ACTTACAGAAACAGTGGCTGACGAGAGAAACATATTGTTGTCGCCCATATCGTTCGACC 1045  
QY 1560 GTGTAAAAAAGCTGTATTGGCTCTCAGACGACCATATTGTGGATGATTAAGAGAAAGTGG 1619  
Db 1047 CTTCAATGACACTTTTGTGCAATGTGACTGATTTCTGTGTAAGAAACACCTCTGCGGTG 1106  
QY 1620 TCCCTCAGGCAATTATACCCACAGCAGACACTGATTTCTGAGGAGCATAGACCGTGG 1679  
Db 1107 GACTGGTGGATAGAGGTAAGAGGACACCGAGATGATTTGGCATATGCTGTGATATT 1166  
QY 1680 CAATACAGATGTGTGGGATTTGTCAAAATTCCTTTGTGGCAGCTCAATGGGCATTCGAG 1739  
Db 1167 GGCTGCCAGGATATGGCTCAGAGGTGGAAGAACTGGGTATTACTGGCTTACACATCAA 1226  
QY 1740 TTCCCTCTTGCACAGCACACACATCAGATTGGACGGCTCAGAGGGGTATGAGTCTAG 1753  
Db 1227 ACACGGGGCACAGGAGGAATAGGACAAAGACTCTCTGACCTGGGGGCCAGTCTGGCGCT 1286  
QY 1800 GCGAGGAATCTTGCAGCTGGAAGCATCTGCTTACTGACTGACTGACACACAGA----- 1850  
Db 1287 CGAGCCCTTGCCTCTGGTTGCTTAAAGGCAACACTTTCCACAGCGAGGCTCGG 1346  
QY 1851 ----CCCTTGGGGGAGTGTCTTCCCATTAATCAGACAGACAGAGGAGTGAATCGCGGA 1907  
Db 1347 GTGGCCACAAGGGGGTGTGTGGAAAAAGATCCCAACTGGAGGCCAGAGAGTGAATTCGGGA 1406  
QY 1908 AAGTACTCTAAAGGGCACACACAGCTGTCTCCGTCACGCTTGGCCATTGCACTGAT 1967  
Db 1407 AAGTACTCTAAAGGGCACACACAGCTGTCTCCGTCACGCTTGGCCATTGCACTGAT 1466  
QY 1968 CCTGGCTTTCGTATGCGGGCCCTCTTGTGGGCATCACCGTCTACTGGCTGTGTGATCA 2027  
Db 1467 CCTGGCTTTCGTATGCGGGCCCTCTCTCTGGGCATCACCGTCTACTGGCTGTGTGATCA 1526  
QY 2028 TCGCGGTAAGAGTGGCTCTGTGGAGGCGCAGGAGAGGAGGCTCACCCACTCGGCGG 2087  
Db 1527 TCGCGGTAAGAGTGGCTCTGTGGAGGCGCAGGAGAGGAGGCTCACCCACTCGGCGG 1586  
QY 2089 GGGTCCATGACGAGGCTCACTAAGCTAGCGGCTCTTTGGGGACATCAATCCAAAGA 2147  
Db 1587 GGGTCCATGACGAGGCTCACTAAGCTAGCGGCTCTTTGGGGACATCAATCCAAAGA 1646  
QY 2148 GCGAAAGCGGAGGCGGCTCTGACGGCACTCACTACAGAGGCAAGCTGCGCACTCGCGG 2207  
Db 1647 GCGAAAGCGGAGGCGGCTCTGACGGCACTCACTGACAGAGGCAAGCTGCGCACTCGCGG 1706  
QY 2208 CAACAGCGGCAAGATGCTTATTAAAGTAGACAGCAGCCACTGTGAGCTGACGGGCTCGG 2257  
Db 1707 CAACAGCGGCAAGATGCTTATTAAAGTAGACAGCAGCAGCAGCTGACGGGCTCGG 1766  
QY 2258 CACCCAGAGTCAACCCCAAGCTGTGAGCAGAGAGGAGGAGCAGCGCGGGAGCGCGGA 2327  
Db 1767 CACCCAGAGTCAACCCCAAGCTGTGAGCAGAGAGGAGGAGGAGCAGCGCGGGAGCGCGGA 1826  
QY 2328 CTGGAGAGAGAACAGAACTCTATCATATGCTTGCACAAAGAGCATAGCCGCGGCTGCTG 2387  
Db 1827 GTGGAGAGGAGAACAGAACTCTATCATATGCTTGCACAAAGAGCATAGCCGCGGCTG 1886  
QY 2388 CCCTGTGATTCGACAGACACTTGGCTGTGGGGGCTGTGCTGAGGAGCAGATCCCGAGCTGG 2447  
Db 1887 CCCTGTGATTCGACAGACACTTGGCTGTGGGGGCTGTGCTGAGGAGCAGATCCCGAGCTGG 1946  
QY 2448 GGTCTCCCATCAGCAGAGAGGCTTACAGCATGACTAGTGGAGCAGCCCAAAATGAG 2507  
Db 1947 GGTCTCCCATCAGCAGAGAGGCTTACAGCATGACTAGTGGAGCAGCCCAAAATGAG 2066  
QY 2508 CGAGGTGGCCAGATGGGCTGGAGGAGCAGGCGGCACACTGGAGTATAAGACCATCAA 2567

Db 2007 CGAGGTGGGCCAGATGGCGCTGGAGGAGCAGGCGGCCACACTGGAGTATATAGACCATCAA 2066  
QY 2568 GGAACATCTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGTGGAGAACTGGACAG 2627  
Db 2067 GGAACATCTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGTGGAGAACTGGACAG 2126  
QY 2628 CTTGCCGCCCAAAAGTTCCACAGCGGAGGCGCTTCCCTGGGTCCCGCGGAGGCTCCCTGTC 2687  
Db 2127 CTTGCCGCCCAAAAGTTCCACAGCGGAGGCGCTTCCCTGGGTCCCGCGGAGGCTCCCTGTC 2186  
QY 2688 TCACAGCGGCTTAACCAAGCGGCTGGAATGCAACCACTCCTCTTCTACGGGGTTGACTA 2747  
Db 2187 TCACAGCGGCTTAACCAAGCGGCTGGAATGCAACCACTCCTCTTCTACGGGGTTGACTA 2246  
QY 2748 TAAGAGAGTACCCACAGAACTCGCTCAGGAGAGGCGGACCCAGGCGCACCTCTCAAAAG 2807  
Db 2247 TAAGAGAGTACCCACAGAACTCGCTCAGGAGAGGCGGACCCAGGCGCACCTCTCAAAAG 2306  
QY 2808 AACAACACTTAACCTCTCCCAATTCCTCTCACCTCTCCAGAAACACAGAGCTTGGCAGGGG 2867  
Db 2307 AACAACACTTAACCTCTCCCAATTCCTCTCACCTCTCCAGAAACACAGAGCTTGGCAGGGG 2366  
QY 2868 AGACAACCCCGCGCGCGCGCGCAGAGGGTGGAGTCCATCCAGGTSCACAGTCCCGACGC 2927  
Db 2367 AGACAACCCCGCGCGCGCGCGCAGAGGGTGGAGTCCATCCAGGTSCACAGTCCCGACGC 2426  
QY 2928 ATCTGCGCAGGCGGTGACTGTCTCGAGGCGAGCGGCTCAAGGCTCAAGGCTCAAGTCACTGAC 2987  
Db 2427 ATCTGCGCAGGCGGTGACTGTCTCGAGGCGAGCGGCTCAAGGCTCAAGGCTCAAGTCACTGAC 2486  
QY 2988 AAGCTCGGGCTTGAAGCGTACGCGCTCGGTAAAGCGGAGGACCGGCTCAAGGCTCAAGGCTCAAGGCT 3047  
Db 2487 AAGCTCGGGCTTGAAGCGTACGCGCTCGGTAAAGCGGAGGCTCAAGGCTCAAGGCTCAAGGCT 2546  
QY 3048 TGCTCCCTTTCCACATCCCATGAAGCCCATGATGCGGTGTACATAA 3093  
Db 2547 TGCTCCCTTTCCACATCCCATGAAGCCCATGATGCGGTGTACATAA 2592

Search completed: September 30, 2003, 16:58:34

Job time : 808.168 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:35:33 ; Search time 5937.36 Seconds  
(without alignments)  
12661.132 Million cell updates/sec

Title: US-09-856-681-1

Perfect score: 3093

Sequence: 1 atgaagtcagaagcttgcttgc.....ccatgatcgtgtacataa 3093

Scoring table: IDENTITY\_NJC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estov:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_estl:\*

10: gb\_estl:\*

11: gb\_hic:\*

12: gb\_estl:\*

13: gb\_estl:\*

14: gb\_estl:\*

15: em\_estin:\*

16: em\_estro:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_hum:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rtd:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrt:\*

28: gb\_gssl:\*

29: gb\_gssl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3089.8	99.9	6875	11	BC032619
2	1439.4	46.5	3226	11	AK042751
3	1437.6	46.5	3329	11	AK082711
4	1065.6	34.5	2411	14	CB605722

5	899.4	29.1	1030	12	BM450002
6	896	29.0	1183	12	BM546059
7	828.2	26.3	568	9	AU140366
8	813.8	26.3	591	13	B0186963
9	784.4	25.4	887	12	BG769297
10	783.4	25.3	1201	9	AL543344
11	779.4	25.2	848	13	BQ678536
12	772.4	25.0	883	13	B0172225
13	770	24.9	864	13	B0440312
14	767.4	24.8	874	13	B0838082
15	760.5	24.6	890	13	B0683009
16	757.4	24.5	953	13	B0854884
17	734	23.7	952	13	B0855855
18	713.8	23.1	835	10	BG477592
19	682.2	22.1	4374	11	AK031307
20	661.8	22.0	3921	11	AK084922
21	680.2	22.0	4476	11	AK052232
22	670.2	21.7	716	10	BF970807
23	667	21.6	890	13	B0856543
24	664.8	21.5	736	10	HE277845
25	660.4	21.4	794	10	AW954605
26	656.2	21.2	823	10	BG327694
27	645.8	20.9	651	9	AL602452
28	636.6	20.6	643	10	HE408781
29	631.8	20.4	685	10	HE265000
30	629.8	20.4	703	10	HE384511
31	615.2	19.9	926	10	BG326467
32	614.4	19.9	746	13	BQ770491
33	614.4	19.9	795	14	CB245479
34	612.2	19.8	772	13	B0425051
35	611.8	19.8	895	12	B1819955
36	608.8	19.7	777	13	B0708565
37	604.4	19.5	771	14	CA513024
38	589.6	19.1	829	12	B1824613
39	589.2	19.0	690	10	BG33712
40	588.8	19.0	592	14	CA337090
41	588.6	19.0	884	10	HE726944
42	578	18.7	603	10	BG334575
43	576.8	18.6	852	10	BF306161
44	573	18.5	713	10	BF389749
45	570.4	18.4	612	10	BG327283

#### ALIGNMENTS

RESULT 1  
BC032619 6875 bp mRNA linear HTC 04-MAR-2003  
LOCUS Homo sapiens, similar to sema domain, transmembrane domain (TM) and cytoplasmic domain. (semaforin) 6A, clone IMAGE:578065, mRNA.  
DEFINITION BC032619  
ACCESSION BC032619  
VERSION BC032619.1 GI:22749800  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: ATCC/DCID/BTP  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC).

Gaithersburg, Maryland:  
 Web Site: <http://www.nisic.nih.gov/>  
 Contact: [nisc.mgcn@nih.gov](mailto:nisc.mgcn@nih.gov)  
 Akhtar, N., Ayello, K., Beckstrom-Stanberg, S.M., Benjamin, B.,  
 Binkley, K.W., Bouffard, G.O., Brown, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,  
 Hansen, N., Ho, L., Karlis, E., Kwon, P., Lari, P., Legasi, R.,  
 Maduro, Q.L., Mastello, C., Masker, B., Magazan, S., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Shankar, S., Thomas, P.J., Touchman, W.,  
 Tusargen, C., Voigt, J.L., Walker, M.A., Welch, K.S., Wilkins, L.,  
 Young, A., Zhang, L.H. and Green, E.P.

Clone distribution: M3C clone distribution information can be found through the I.M.A.G.E. Consortium/LLN at: <http://image.lln.gov>  
Series: IRAK Plate: 69 Row: 0 Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA at: 11991459  
This clone has the following problem: retained intron:

FEATURES  
SOUND

```

1784 a 1795 c 1542 z 1704 t
/organism="Homo sapiens"
/mol_type="mrna"
/cdb_xref="taxon:9606"
/closure_image="5578066"
/closure_type="skin, melanotic melanoma."
/closure_lib="NH1MG-72"
/cdb_host="DH10B"
/clone_vector="pCMV-Sport6"

```

Query Match 99.9%; Score 3089.8; DB 11; Length 6875;  
Best Local Similarity 99.9%; Prod. No. 0;  
Matches 3091; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGAGGCTCAGAAGCCTTGCCTGTATATTTCCACACATGCTACACATTTCGTGGGGCTGGTTC	60
Db	752	ATGAGGCTCAGAAGCCTTGCCTGTATATTTCCACACATGCTACACATTTCGTGGGGCTGGTTC	851
QY	61	CCAGAAGATTCAGCCAAATCAGATTTCCCATGGCCAAATATACAAAACACATATCCCGGTG	129
Db	852	CCAGAAGATTCAGCCAAATCAGATTTCCCATGGCCAAATATACAAAACACATATCCCGGTG	911
QY	132	TTTCTGGGCTCACAAGCCAGACGGSACACACACACAGACGACAGCTGGACATCCAGATG	180
Db	912	TTTCTGGGCTCACAAGCCAGACGGSACACACACACAGACGACAGCTGGACATCCAGATG	973
QY	181	ATTATGATCATGAACGACGACCTACATTTGCTGTAGGACCATATTTATACGTGTGAT	240
Db	972	ATTATGATCATGAACGACGACCTACATTTGCTGTAGGACCATATTTATACGTGTGAT	1031
QY	241	ATAGACATCATCACACCGGAGAAATTTATTTAGCAAAAACACTGCATGSAATCTAGA	300
Db	1032	ATAGACATCATCACACCGGAGAAATTTATTTAGCAAAAACACTGCATGSAATCTAGA	1091
QY	301	CAGGCCGATGTAGACATATGCAATATGAGGAGAAACATAGCATGAGCTGCCACATTT	360
Db	1092	CAGGCCGATGTAGACATATGCAATATGAGGAGAAACATAGCATGAGCTGCCACATTT	1151
QY	361	ATTAAGATTTCTTAAGAAAAACGATGATGCATTTTGTGTGTSAACTAATGCTTC	420
Db	1152	ATTAAGATTTCTTAAGAAAAACGATGATGCTTTTGTGTGTSAACTAATGCTTC	1211
QY	421	AAGCCTTCTCCAGAAACTATAGATGGATACATATGGAAACATTCGGGATCAATTCAGC	480
Db	1212	AAGCCTTCTCCAGAAACTATAGATGGATACATATGGAAACATTCGGGATCAATTCAGC	1271
QY	481	GGAAATGGCCAGATGCCCATATGATGCCAAACATGCAAGCTTGCACTTTTGCAGATGGA	540
Db	1272	GGAAATGGCCAGATGCCCATATGATGCCAAACATGCAAGCTTGCACTTTTGCAGATGGA	1331
QY	541	AAACTATATCTCAGCCACAGTGGACTTCCTTGGCATATGAGCGGACTCATTTACCGGAGT	600



Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasole, G., Quackenbush, J., Schriml, M., Staebli, F., Suzuki, K., Tomita, M., Wagner, I., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balcells, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Rinaldo, M., Brownstein, M., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustigich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Marzari, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, R., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S., and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 695-698 (2001)

21045660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II team

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3226)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, P., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiracka, T., Hirozane, T., Hori, F., Inotani, K., Isuli, Y., Iton, M., Kageoka, I., Kasakawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, K., Koya, S., Kurikawa, C., Matsuyama, I., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nomazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saito, H., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki, A., Takada, Y., Tanaka, T., Tomaru, A., Toya, I., Yasunishi, A., Yamamatsu, M., and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://phantom.gsc.riken.go.jp/

Location/Qualifiers

1..3226

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM\_DB:A73020P05"

/db\_xref="taxon:10090"

/clone="A73020P05"

/tissue\_type="cerebellum"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="7 days neonate"

671..2488

/note="unnamed protein product; putative scd4 domain, transmembrane domain (TM), and cytoplasmic domain, (senaphorin) 5A (MGI:MG:1263727, GRIN:018744, evidence: BLASTN, 100%, match=3802)"

/codon\_start=1

/protein\_id="BAC31351.1"

/db\_xref="gb:24335301"

/translation="KRPAALLLCLTLHCAGAGFPEDSEPTISIHGNTKQVPFVGH KPGNTQTQRHLDIQIMIMNNTLYVAARDHIYIVDIDISHTEEIYQSKSLTKWSRQA

DVIDICPMKCKHKHDECHNHFKVLLKKNDOTLFVCGTNAENPSCRNVRVDTLFTFGDEFS  
CHACPDYDAKHANIALPADCKLYSATVDFDLAIDAVYPSLCDSTPLRVHDKWLK  
EPFVQVVDIGSTIFFPREIAVEINTKGVVFPVQVCKNDGMSORVLEKQWTS  
LKLRLNCVPGSHFYFNILQAVTDIRKNGVAVLATFSPINSPUSACVDMUD  
IANVTGREKFKQSPDSTWTPDPRVPRKPPGCCAGSSSLEKATSNFEPDITLNF  
KTIPLIMDEAVPSTINRPMFLRTMVAKYRLTKIAVDNAGPYQNHITVPLGSEKGLIKF  
LARISSGFLNGLSFLFEMNVTNPKCSYDGVEDKRIKINGMOLDRAKSGSLYVAFSTCVI  
KVLPSRCEHCKKCTCIASRDYPCWVRESGCAHLSPLSRLTFEODIERGNTDGLG  
KPLSPCEHCKKCTCIASRDYPCWVRESGCAHLSPLSRLTFEODIERGNTDGLG  
DCHNSEVALNDISITPLPDYKSEFNRYGFIATIIYIN"

BASE COUNT 883 a 752 c 731 g 860 t

CRISIN

Query Match

Best Local Similarity 46.5%; Score 1439.4; DB 11; Length 3226;

Matches 1554; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Q/ 1 ATGAGTCCAGAGCCCTGCTGTATATTTTACACTGCTACACTTTGCTGGGCTGGTTTC 60

DE 671 ATGGGCCACAGCCTTACTGCTGTCTCACACTGCTACACTGGCTGGGCTGGTTTC 730

Q/ 61 CCAGAAATCTCTGAGCAATCACTATTTCGCATGCGCACTATACAAACAGTATCGGTG 120

DB 731 CCAGAAATCTCTGAGCAATCACTATTTCGCATGCGCACTATACAAACAGTATCGGTG 790

Q/ 121 TTGTGGGCAAGCCAGGACGACACACAGAGGACAGGCTGCACATCCAGATG 180

DB 791 TTGTGGGCAAGCCAGGACGACACACAGAGGACAGGCTGCACATCCAGATG 850

Q/ 181 ATTATGATGATGACGAGGACGCTCTACATGCTGTAGTAGGACATATTTACTGTGAT 240

DB 851 ATCATATGATGACGAGGACGCTCTACATGCTGTAGTAGGACATATTTACTGTGAT 910

Q/ 241 ATAGACACATTCACACGAGGAAATTTATGTAGCAAAAACATGACATGGAATCTAGA 300

DB 911 ATAGACACATTCACACGAGGAAATTTACTGTAGCAAAAACATGACATGGAATCTAGA 970

Q/ 301 CAGGCGATGTAACACATGACAGATGAGAGGAAACATGAAGATGAGTGCACACACTTT 360

DB 971 CAGGCGATGTAACACATGACAGATGAGAGGAAACATGAAGATGAGTGCACACACTTT 1030

Q/ 361 ATTAAGTTCTTCTAAGAAACAGTATGATGCTTTGCTGTGGAACATATGCTTC 420

DB 1031 ATTAAGTTCTTCTAAGAAACAGTATGATGCTTTGCTGTGGAACATATGCTTC 1090

Q/ 421 AACCTCTCCGAGAACTATAAGATGGATACATTTGGAACCATTCGGGATGAATTCAGC 480

DB 1091 AACCTCTCCGAGAACTATAAGATGGATACATTTGGAACCATTCGGGATGAATTCAGC 1150

Q/ 481 GGAATGCCAGATGCCCATATGATGCCCAACATGCCAAGTTCACACTGTTSCAGATGA 540

DB 1151 GGAATGCCAGATGCCCATATGATGCCCAACATGCCAAGTTCACACTGTTSCAGATGA 1210

Q/ 541 AAATCTACTCAGCCACAGTACTGCTTCITGCCATTCAGCAGTCAATTTACCGAGCT 600

DB 1211 AAATCTACTCAGCCACAGTACTGCTTCITGCCATTCAGCAGTCAATTTACCGAGCT 1270

Q/ 601 CTGGAGAAAGCCCTACCTGCGGAGCGCTGACACAGATCAAAATGTTTGAAGAACCA 660

DB 1271 CTGGAGAAAGCCCTACCTGCGGAGCGCTGACACAGATCAAAATGTTTGAAGAACCA 1330

Q/ 661 TACTTTCTTCAGCCCTGGATACGGAGATATATCTACTTCTTCAGGGAATAGCA 720

DB 1331 TACTTTCTTCAGCCCTGGATACGGAGATATATCTACTTCTTCAGGGAATAGCA 1390

Q/ 721 GTGAGTATACACCATGGGAAGGTAGTTTCCCAAGAGTGGCTCAGTGTGTTGAAGAT 780

DB 1391 GTGAGTATACACCATGGGAAGGTAGTTTCCCAAGAGTGGCTCAGTGTGTTGAAGAT 1450

Q/ 781 GATATGGAGGATCTCAAAGAGTCCCTGGGAAACACAGTGGACCTGTTCTCTGAAGCGCGC 840

DB 1451 GATATGGAGGATCTCAAAGAGTCCCTGGGAAACACAGTGGACCTGTTCTCTGAAGCGCGC 1510

Q/ 841 TTGAATGCTCAGTCTCTCGAGAGCTCTCATTTTTTATTTTCAACATTCCTCCAGGACGATTA 900













```

QY 1686 AGAT-GATCGGGGAGTCTGAC--AAATCTTTTGGGACTGAA---TGGGATCTCCAG 1739
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 AGATGGGCTGGGGGAATG/CACCAATCTCTTTGTTGGGCACTGNAATGGGCAATTCAG 960
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1740 TTGCTCTTCCCGAGCACAACC 1761
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 TTGCTCTTCCCGAGCACAACC 982
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
BM546059 1183 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6497880 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5588479
5', mRNA sequence.
ACCESSION BM546059
VERSION 1
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1183)
TITLE N7H-MGC http://mgc.ncl.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM-2359 row: k column: 06
High quality sequence stop: 725.
location/Qualifiers
source 1..1183
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5588479"
/lab_host="DH10h"
/clone_lib="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
Site1: EcoRV (destroyed); Site2: NotI; RNA source pool
of three ovaries, from females ranging in age from 18 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by T. Gruber (Invitrogen). Research: Genomics
tracking code 036."
BASE COUNT 264 a 428 c 309 g 161 t
ORIGIN
Query Match 29.0%; Score 896; DH 127 Length 1183;
Best Local Similarity 99.0%; Pred. No. 1,Je-225;
Matches 932; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 2152 AAGCCGAGGCGCATCTCTACACCACTGTCACCAAGGCAAGCTCGCTACTCGGCGCAAC 2211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18 AAGCCGAGGCGCATCTCTACACCACTGTCACCAAGGCAAGCTCGCTACTCGGCGCAAC 77
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2212 ACGGCCAAGATGTCATTAAAGCAGACAGCACCACCTGGACCTGACGGCGCTCCGCCAC 2271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78 ACGGCCAAGATGTCATTAAAGCAGACAGCACCACCTGGACCTGACGGCGCTCCGCCAC 137
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2272 CCGAGGTCAACCCCAACGCTCCAGAGAGCGGAGCCAGCCGCGGCGGCGGCGGAGTGG 2331
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 CCGAGGTCAACCCCAACGCTCCAGAGAGCGGAGCCAGCCGCGGCGGCGGCGGAGTGG 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2332 GAGAGGAACCCAGACCTCATCTGCTCCACCAAGGACATGCCCCCGCTGGCTCCCT 2391
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 7
BM546059 958 bp mRNA linear EST 05-AUG-2002
AGENCOURT_6497880 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:2000407 5', mRNA
5', mRNA sequence.
ACCESSION BM546059
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 958)
TITLE N7H-MGC http://mgc.ncl.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Takao Isogai
Genomics Laboratory

```

```

RESULT 8
BU186963
LOCUS       BU186963
DEFINITION AGENCOURT_7785723 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6138529
            5', mRNA sequence.
ACCESSION   BU186963
VERSION     BU186963.1 GI:22700947
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 891)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-rt@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13454 Row: b Column: 02
            High quality sequence stop: 677.
            Location/Qualifiers
                1..89;
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6138529"
                /tissue_type="retinoblastoma"
                /lab_hosts="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_67"
                /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.75 kb. Library constructed by Life
                Technologies."
            BASE COUNT      235 a 210 c 226 g 219 t      1 others
            ORIGIN
                Query Match      26.3%; Score 913.8; DB 13; Length 891;
                Best Local Similarity 97.3%; Pred. No. 6.5e-204;
                Matches 860; Conservative 0; Mismatches 18; Indels 6; Gaps 3
                Q7      946  TCTACACTATACAGCATCCCTGGGTGCGAGTGTGGCTATACAGATGTTGACATT 1005
                Db      8   TCGGACCTTATACAGCATCCCTGGGTGCGAGTGTGGCTATACAGATGTTGACATT 67
                QY      1006  GCCAGTGTGTTTACTGGGAGATTCAAGGAACAGAAGTCTCTTGATTCACCTGGACACCA 1065
                Db      68   GCCAGTGTGTTTACTGGGAGATTCAAGGAACAGAAGTCTCTTGATTCACCTGGACACCA 127
                QY      1066  GTTCTGTGATGACAGAGTTCCTTAAGCCAGGCCAGGTGTGCTGTGGCTATCCCTCCCTTA 1125
                Db      128  GTTCTGTGATGACAGAGTTCCTTAAGCCAGGCCAGGTGTGCTGTGGCTATCCCTCCCTTA 187
                QY      1126  CAAAGATATCAACCTCCCAATGAGTTCCTCTGATGATACCTTGAACTTCATCAAGACGCAC 1185

```



Db 662 CATCAAGAACATCTCAAGCAAGAGTCGCAACCAATGGGGTGAACCTGTGCAAGACC 721

Qy 2621 TGGACAGCTGCCCCCAAGATTTCACAGCGGGAGCGCTCCCTGGTCCCGCGGGA--GC 2678  
|||||

Db 722 TGGACAGCTGCCCCCAAGATTTCACAGCGGGAGCGCTCCCTGGTCCCGGTAGCC 781

Qy 2679 CTCCTGTCTCAGACGGGCTTAAGCAAGCGCTGGAAATGACACACACCTCTCTCTACGG 2718  
|||||

Db 782 TCCCTGTCTTACAGCGGTTCAGCAAGCGCTGAAATGACACACCTCTCTCTCTACGG 841  
|||||

Qy 2739 GCTT--GACTATAAGAGAGCTACCCACGAACTGCTCAGAGAA 2782  
|||||

Db 842 GGTAGCACTTATAGAGCACTTAACCAAGCAAGTCTGCTCAAGACA 887

RESULT 10  
AL543344

LOCUS AL543344 Homo sapiens PLACENTA.COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CSOD1001YL17 5-PRIME, mRNA sequence.

ACCESSION AL543344

VERSION AL543344.2 GI:31255191

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
Li, W.B., Gruber, C., Jessoe, J., and Pelley, D.  
Full-length cDNA libraries and normalization  
Unpublished  
On Feb 15, 2001 this sequence version replaced gi:12875822.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by life technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 2864.r for  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOD1001CF090PL&cl=ster-2864.r. Contact :  
Peng Liang Email : liliang@life.techn.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSOD1001CF090PL  
Location/Qualifiers  
1..120:  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1001YL17"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with NotI and cloned into the NotI and EcoRV  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 277 a 373 c 315 g 214 t  
ORIGIN

Query Match 25.3% Score 783.4; DB 9; Length 1201;  
Best Local Similarity 96.4%; Pred No. 8.8e-196;  
Matches 838; Conservative 5; Mismatches 2; Indels 5; Gaps 4;

Qy 1714 TTGTGGCACTGAATGGGATTCACATTCCTGCTTTGGGCAACCAACCAATCAGATTCG 1773  
|||||

Db 221 TTGTGATGAAGAGGGATTCACATTCCTGCTTTGGGCAACCAACCAATCAGATTCG 280  
|||||

Qy 1774 AGCGTCAAGAGGGATTCACATTCCTGCTTTGGGCAACCAACCAATCAGATTCG 1833  
|||||

Db 281 AGCGTCAAGAGGGATTCACATTCCTGCTTTGGGCAACCAACCAATCAGATTCG 349  
|||||

Qy 1834 TCACCTGACAGCAGACACCTTTGGGCGAGTGTCTTCCCATATATCAACCAAGCAAG 1893  
|||||

Db 341 TCACCTGACACACAGAACCCCTTTGGGGGAGTGTCTTCCCATATATCAACCAAGCAAG 400

Qy 1894 GGAGTGATTGGGGAAGTTACCTCAAGAGCCAGACAGCTGGTTCCGCTCACCTCTTG 1953  
|||||

Db 401 GGAGTGATTGGGGAAGTTACCTCAAGAGCCAGACAGCTGGTTCCGCTCACCTCTTG 460

Qy 1954 GCATTTGAGTTCATCTCTGCTTATGGGGGCGCTTCTTCGCGGCATCACCTCTAC 2013  
|||||

Db 461 GCATTTGAGTTCATCTCTGCTTATGGGGGCGCTTCTTCGCGGCATCACCTCTAC 520

Qy 2014 TCGGCTGTGATCATCGGCGAAGAGCTGTGTGTGCGAGCGCAAGGAGAGAGCTC 2073  
|||||

Db 521 TCGGCTGTGATCATCGGCGAAGAGCTGTGTGTGCGAGCGCAAGGAGAGAGCTC 580

Qy 2074 ACCGACTCGGCGGCGCTTCATGAGCAGCGCTCAACAGCTCAGCGGCTCTTTGGG 2133  
|||||

Db 581 ACCGACTCGGCGGCGCTTCATGAGCAGCGCTCAACAGCTCAGCGGCTCTTTGGG 640

Qy 2134 ACTCAATCAAGAGCCCAAGCGCGAGCGCTTCAGCGGCATCTCAGCGGCATCTGCAACAGCGCAAG 2193  
|||||

Db 641 ACTCAATCAAGAGCCCAAGCGCGAGCG--CATCTTCAGCGGCATCTGCAACAGCGCAAG 699

Qy 2194 CTGCGCACTCCCGGCAACAGCGCGAGATGTCTATTAAAGCAGACCGACACACCTGGAC 2253  
|||||

Db 700 CTGCGCACTCCCGGCAACAGCGCGAGATGTCTATTAAAGCAGACCGACACCTGGAC 759

Qy 2254 CTGAGCGGCTTCACACCGCAGAGTCAACCGCAGCGCTGCAGCAGAGCGGAGCGGACG 2313  
|||||

Db 760 CTGAGCGGCTTCACACCGCAGAGTCAACCGCAGCGCTGCAGCAGAGCGGAGCGGACG 819

Qy 2314 CCGCGCAGCGCGGATTTGGAGAGAGCAACAGACCTCATCAATGGCTGCACAAAGGACATG 2373  
|||||

Db 320 CCGCGCAGCGCGGATTTGGAGAGAGCAACAGACCTCATCAATGGCTGCACAAAGGACATG 879

Qy 2374 CCGCGCAGCGCGCTTCCTGTGATTCACAGGACCTTGGCGCTTGGCGGCTTCCCGCAGCGAC 2433  
|||||

Db 880 CCGCGCAGCGCGCTTCCTGTGATTCACAGGACCTTGGCGCTTGGCGGCTTCCCGCAGCGAC 939

Qy 2434 ATCCGAGCGGTGTGCTCTGCCATCAGCGCAGCGCTACCGACATGAGTACGTGGAC 2493  
|||||

Db 940 ATCCGAGCGGTGTGCTCTGCCATCAGCGCAGCGCTACCGACATGAGTACGTGGAC 999

Qy 2494 CAGCCCAATAGCGAGGTGCGCCAGATGCGCGCTGTGGAGGAGCGGCGCCACATGGAG 2553  
|||||

Db 1000 --AGCCCAATAGCGAGGTG--CCAGATGGCGTGGAGGCCAGG--CGSCACACTGGAG 1055

Qy 2554 TATAGACCATCAGAGACATCTCAGCAG 2582  
|||||

Db 1056 TATAGACCATCAGAGAAATYAGCAGCAG 1084  
|||||

RESULT 11  
B0678536

LOCUS B0678536 848 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGENCOURT\_8208159 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6262572  
5', mRNA sequence.

ACCESSION B0678536

VERSION B0678536.1 GI:21791215

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 848)  
NIH-MGC http://imgc.nhl.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: rgs@bbs.nhl.nih.gov  
Tissue Procurement: DCTD/PTP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation





```

QY 2316 CGCAGCGCGG-AGTGGGAGGAGACA-GAAGCTCATCAATGCGGACGAAGAGACATG 2373
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 CGCAACCGCGGAGTGGGAGAACACANNAAGCTCATCAATGCGGACGAAGAGACATG 780
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2374 CCCCCCAATGGG-CTCCCTCTGTGAATCCACAGGAGCTG-----CCCTGGGGGCTCCCG 2427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 CCCCCCAATGGGGCTCCCTCTGTGAATCCACAGGAGCTG-----CCCTGGGGGCTCCCG 840
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2428 AGCCACATCCCGACGCGTGGG 2459
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 AGCACATCCCGCCCAAGCGGG 862
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
B0838082
LOCUS B0838082 874 bp mRNA linear EST 15-OCT-2002
DEFINITION AGENCOURT_8071571 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088226
5' mRNA sequence.
ACCESSION B0838082
VERSION B0838082.1 GI:24022477
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 874)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cyapbs-remail.nih.gov
Tissue Procurement: DfID/DFP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L2CM236 row: b column: 03
High quality sequence stop: 652.
Location/Qualifiers
1..874
/organism="Homo sap-ens"
/molecule-type="cDNA"
/db_xref="Genbank:9605"
/clone="IMAGE:608224"
/rdb_host="G410B (phage-resistant)"
/rdb_type="melanoma, cell: line"
/clone_lib="NIH_MGC_112"
/notes="Organ: Skin; Vector: pCIS7; Site: 2; XhoI: Site 2;
Note: cDNA made by oligo-dT priming. Directionally cloned
into EcoR/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."
BASE COUNT 213 a 211 c 247 g 203 t
ORIGIN
Query Match
Best Local Similarity 99.2% Score 767.4; DB 13; Length 874;
Matches 771; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1286 CTGGGCGCATATCAGATCACTGCGTTTCTGGGATCAGAGAGGAATCATCTTGA 1345
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CTGGGCGCATATCAGATCACTGCGTTTCTGGGATCAGAGAGGAATCATCTTGA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1346 AGTTTTCGGCAGATAGGAATAGTGGTTTCTAAATGACACCTTTCTCGAGGAGA 1405
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AGTTTTCGGCAGATAGGAATAGTGGTTTCTAAATGACACCTTTCTCGAGGAGA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 1406 TGAATGTTTAAACACTCTGAATAATGCAGCTATGATGAGTGCAGACAAAAGGATCATGG 1465
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TGAATGTTTAAACACTCTGAATAATGCAGCTATGATGAGTGCAGACAAAAGGATCATGG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1466 GCATGCACTGGACAGACAGACAGCTCTGTATGTTGCGTCTCTACCTGTTGATAA 1526
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GCATGCACTGGACAGACAGACAGCTCTGTATGTTGCGTCTCTACCTGTTGATAA 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1526 AGTTTTCCTCTGGCGGTGTGAAGCAGATGGAGTGTAAATAAACCTGTATTGCCTCCA 1585
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGTTTTCCTCTGGCGGTGTGAAGCAGATGGAGTGTAAATAAACCTGTATTGCCTCCA 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1586 GAGACCATATGTGATGGATAAAGAGGTGGCTGCGAGCATTTATACCCCAACA 1645
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GAGACCATATGTGATGGATAAAGAGGTGGCTGCGAGCATTTATACCCCAACA 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1646 CGAGACTCATTGAGCAGGACATAGAGCTGCAATACAGATGGTCTGGGGACTGTC 1705
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CGAGACTCATTGAGCAGGACATAGAGCTGCAATACAGATGGTCTGGGGACTGTC 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1706 ACAATTTCCTTTGTGCACTGAATGGGCAATCCAGTTCCCTCTGCCAGCACAACCAT 1765
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ACAATTTCCTTTGTGCACTGAATGGGCAATCCAGTTCCCTCTGCCAGCACAACCAT 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1766 CAGATTGACGCCICAGAGGGGTATGAGTCTAGGGGAGGAATGCTGGACTGGAGCATC 1825
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 CAGATTGACGCCICAGAGGGGTATGAGTCTAGGGGAGGAATGCTGGACTGGAGCATC 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1826 TCGTTGACTTCACCTGACAGCAGACACCCCTTTGGGGGAGTGTCTTCCCATATCACCAG 1885
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 TCGTTGACTTCACCTGACAGCAGACACCCCTTTGGGGGAGTGTCTTCCCATATCACCAG 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1886 ACAGAAGGGAGTGATTGGGAAAGTTACCTCAAGGCCAGACAGAGCTGGTCTCCGACA 1945
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 ACAGAAGGGAGTGATTGGGAAAGTTACCTCAAGGCCAGACAGAGCTGGTCTCCGACA 660
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1946 CCTCTTGGCAATTCAGTATCCTGCTTTCGTCATGCGGGCGCTCTCTCGGGCATCA 2005
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 CCTCTTGGCAATTCAGTATCCTGCTTTCGTCATGCGGGCGCTCTCTCGGGCATCA 720
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2006 CCGTTACTGCTGTGTATCATCATCGGCCAAGAGCGCTGTGTGTCAGGCGCAAG 2062
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 CCGTTACTGCTGTGTATCATCATCGGCCAAGAGCGCTGTGTGTCAGGCGCAAG 777
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RFSL15
B0683009
LOCUS B0683009 590 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_820R014 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262587
5' mRNA sequence.
ACCESSION B0683009
VERSION B0683009.1 GI:21795688
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 890)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cyapbs-remail.nih.gov
Tissue Procurement: DfID/DFP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L2CM2425 row: c column: 04
High quality sequence stop: 621.
Location/Qualifiers

```

FEATURES



```
source 1..890
/organism="Homo sapiens"
/genetype="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6262587"
/issue="melanotic melanoma, cell line"
/lab_host="BHI0B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/notes="organ: skin; Vector: pORF7; Size:1; XhoI: Site:2;
ECORI: cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."
```

```
220 a 220 c 245 g 204 t
721 CCGTCTACTGGGCTCTGTGATCATCGCGCAAAAGACGTGGCTGSSGGCACCAGCGA 780
QY 2063 --AGAGAGAGGCTCACCCTACTCCCGCGGGGGCTCCCA-TCAGCAGGCTCACCACAG 2112
DB 781 GAAGGAGCTTCACCCATTCCTCCCGCGGGGGGCTCCCATTCAGAGGGGCTCACCACAG 833
```

Search completed: September 30, 2003, 22:02:09  
Job time : 5947.36 secs

```
BASE COUNT 220 a 220 c 245 g 204 t
ORIGIN

Query Match 24.6%; Score 760.6; E: 13; Length 890;
Best Local Similarity 97.0%; Pred. No. 46-150;
Matches 808; Conservative (-); Mismatches 19; Indels 6; Gaps 7;

QY 1286 CTGGGCGCATCATCAGATCATCTGTGGTTTCTGTGATCAGAGAGGAGATCATCTGTA 1345
DB 1 CTGGGCGCATCATCAGATCATCTGTGGTTTCTGTGATCAGAGAGGAGATCATCTGTA 80

QY 1346 AGTTTGGCCAGATAGGAATAGTGGCTTTCTTAAAGATGAGAGGCTTTCTGTGAGAGACA 1405
DB 61 AGTTTGGCCAGATAGGAATAGTGGCTTTCTTAAAGATGAGAGGCTTTCTGTGAGAGACA 120

QY 1406 TGAGTGTTCACACTGGAAGATGAGAGATGATGATGAGAGATGAGAGAGAGATGATGAG 1465
DB 121 TGAGTGTTCACACTGGAAGATGAGAGATGATGATGAGAGATGAGAGAGAGATGATGAG 180

QY 1466 GCATGACGCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1525
DB 181 GCATGACGCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240

QY 1526 AGTTCCCTCTGGCGCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1585
DB 241 AGTTCCCTCTGGCGCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 1586 GAGACCCATATTGTGGATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645
DB 301 GAGACCCATATTGTGGATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 1646 GCAGACTGACTTTTGACAGGAGATAGAGGCTGGCAATACAGATGGTCTGGGGAGCTGTC 1705
DB 361 GCAGACTGACTTTTGACAGGAGATAGAGGCTGGCAATACAGATGGTCTGGGGAGCTGTC 420

QY 1706 ACAATTCCTTTTGGCAGCTGAATGGGCAATCCAGATTCCTCTGGCAGCAGACACACAT 1765
DB 421 ACAATTCCTTTTGGCAGCTGAATGGGCAATCCAGATTCCTCTGGCAGCAGACACACAT 480

QY 1766 CAGATTCGAGGGCTCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1825
DB 481 CAGATTCGAGGGCTCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

QY 1826 TGCTTGACTCAGCTGACAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1885
DB 541 TGCTTGACTCAGCTGACAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

QY 1886 ACAAGAGGGAGTGAATTCGGGAAAGTTACCTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1945
DB 601 ACAAGAGGGAGTGAATTCGGGAAAGTTACCTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 1946 CCCTCTTGGCCATTCGACATGCTGGCTTCTGCTCAAGGGGGGGTCTCTCTGGGGATCA 2005
DB 661 CCCTCTTGGCCATTCGACATGCTGGCTTCTGCTCAAGGGGGGGTCTCTCTGGGGATCA 720

QY 2006 CCGTCTACTGGCTGTGATCATCGGCGCAAG-AGGTGGCTGTGGTGGAGGCGCAAGG-- 2062
DB 720 CCGTCTACTGGCTGTGATCATCGGCGCAAG-AGGTGGCTGTGGTGGAGGCGCAAGG-- 2062
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003: Comptagen Inc.

OM nucleic acid - nucleic search, using sw mode.

Run on: September 30, 2003, 16:44:06 : Search time 761.8 Seconds  
(without alignments)  
10107.639 Million cell updates/sec

Title: US-09-856-681-1

Perfect score: 3093

Sequence: 1 atgaggagcagaagccttgcct.....ccaatgatgaggtatataaa 3293

Scoring table: IDENTITY\_NUC

Gapop 10-0 , Gapext 1-0

Searched: 1678620 seqs, 1244745471 residues

Total number of hits satisfying chosen parameters: 3357240

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 200%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_5/ptodata/2/pubpna/US07\_PUBSCOMB.seq.\*  
2: /cgn2\_5/ptodata/2/pubpna/PC1\_NEW\_PUB.seq.\*  
3: /cgn2\_5/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_5/ptodata/2/pubpna/US06\_PUBSCOMB.seq.\*  
5: /cgn2\_5/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_5/ptodata/2/pubpna/PC1US\_PUBSCOMB.seq.\*  
7: /cgn2\_5/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_5/ptodata/2/pubpna/US08\_PUBSCOMB.seq.\*  
9: /cgn2\_5/ptodata/2/pubpna/US09A\_PUBSCOMB.seq.\*  
10: /cgn2\_5/ptodata/2/pubpna/US09B\_PUBSCOMB.seq.\*  
11: /cgn2\_5/ptodata/2/pubpna/US09C\_PUBSCOMB.seq.\*  
12: /cgn2\_5/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_5/ptodata/2/pubpna/US10A\_PUBSCOMB.seq.\*  
14: /cgn2\_5/ptodata/2/pubpna/US10B\_PUBSCOMB.seq.\*  
15: /cgn2\_5/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_5/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_5/ptodata/2/pubpna/US60\_PUBSCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3038	98.2	3498	11	US-09-991-053-3 Sequence 3, Appl
2	3036	98.2	3498	11	US-09-957-187-3 Sequence 3, Appl
3	3030.4	98.0	4250	11	US-09-957-187-84 Sequence 84, Appl
4	2698	87.2	3333	11	US-09-991-053-5 Sequence 5, Appl
5	2668	87.2	3333	11	US-09-957-187-5 Sequence 5, Appl
6	1981.6	60.8	1890	11	US-09-991-053-29 Sequence 29, Appl
7	1881.6	60.8	1890	11	US-09-957-187-29 Sequence 29, Appl
8	1881.6	60.8	1890	11	US-09-957-187-82 Sequence 82, Appl
9	604	19.5	845	10	US-09-933-381-920 Sequence 920, Appl
10	565	18.3	1472	9	US-09-925-299-66 Sequence 66, Appl
11	565	18.3	1472	11	US-09-925-299-60 Sequence 60, Appl
12	555.4	18.0	3524	12	US-10-391-413-3 Sequence 3, Appl
13	553.8	17.9	3721	11	US-09-933-836-34 Sequence 34, Appl
14	553.8	17.9	3721	12	US-10-035-977-34 Sequence 34, Appl
15	553.8	17.9	3721	12	US-10-137-870-543 Sequence 543, Appl
16	553.8	17.9	3721	12	US-10-140-018-543 Sequence 543, Appl

17	553.8	17.9	3721	12	US-10-140-021-543	Sequence 543, App
18	553.8	17.9	3721	12	US-10-140-274-543	Sequence 543, App
19	553.8	17.9	3721	12	US-10-140-471-543	Sequence 543, App
20	553.8	17.9	3721	12	US-10-140-807-543	Sequence 543, App
21	553.8	17.9	3721	12	US-10-140-922-543	Sequence 543, App
22	553.8	17.9	3721	12	US-10-140-924-543	Sequence 543, App
23	553.8	17.9	3721	12	US-10-140-925-543	Sequence 543, App
24	553.8	17.9	3721	12	US-10-141-668-543	Sequence 543, App
25	553.8	17.9	3721	12	US-10-141-704-543	Sequence 543, App
26	553.8	17.9	3721	12	US-10-141-704-543	Sequence 543, App
27	553.8	17.9	3721	12	US-10-142-421-543	Sequence 543, App
28	553.8	17.9	3721	12	US-10-142-432-543	Sequence 543, App
29	553.8	17.9	3721	12	US-10-142-767-543	Sequence 543, App
30	553.8	17.9	3721	12	US-10-143-033-543	Sequence 543, App
31	553.8	17.9	3721	12	US-10-144-994-543	Sequence 543, App
32	553.8	17.9	3721	12	US-10-145-628-543	Sequence 543, App
33	553.8	17.9	3721	12	US-10-145-631-543	Sequence 543, App
34	553.8	17.9	3721	12	US-10-145-633-543	Sequence 543, App
35	553.8	17.9	3721	12	US-10-145-746-543	Sequence 543, App
36	553.8	17.9	3721	12	US-10-145-748-543	Sequence 543, App
37	553.8	17.9	3721	12	US-10-145-823-543	Sequence 543, App
38	553.8	17.9	3721	12	US-10-145-836-543	Sequence 543, App
39	553.8	17.9	3721	12	US-10-145-870-543	Sequence 543, App
40	553.8	17.9	3721	12	US-10-145-876-543	Sequence 543, App
41	553.8	17.9	3721	12	US-10-145-959-543	Sequence 543, App
42	553.8	17.9	3721	12	US-10-146-724-543	Sequence 543, App
43	553.8	17.9	3721	12	US-10-146-725-543	Sequence 543, App
44	553.8	17.9	3721	12	US-10-146-795-543	Sequence 543, App
45	553.8	17.9	3721	12	US-10-147-495-543	Sequence 543, App

#### ALIGNMENTS

RESULT :  
US-09-991-053-3  
: Sequence 3, Application US/09991053  
: Publication No. US2003003532A1  
: GENERAL INFORMATION:  
: APPLICANT: Shimkots, Richard A.  
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES ENCODING HUMAN SLIT-,  
: FILE OF INVENTION: MEGF-, AND ROUNDABOUT-LIKE POLYPEPTIDES  
: FILE REFERENCE: 15966-540 CON S-10  
: CURRENT APPLICATION NUMBER: US/09/991, 053  
: CURRENT FILING DATE: 2002-05-23  
: PRIOR APPLICATION NUMBER: USSN 60/123,667  
: PRIOR FILING DATE: 1999-03-09  
: PRIOR APPLICATION NUMBER: 09/520,781  
: PRIOR FILING DATE: 2000-03-08  
: NUMBER OF SEQ ID NOS: 81  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO: 3  
: LENGTH: 3498  
: TYPE: CNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (214)...(3630)  
: FEATURES:  
: NAME/KEY: misc\_feature  
: LOCATION: (3047)  
: OTHER INFORMATION: an s may be any one of a or t or q or c  
US-09-991-053-3

Query Match: 98.2%; Score 3038; DB 11; Length 3498;

Best Local Similarity: 99.3%; Pred. No. 0;

Matches 3072; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

QY 1 ATCAGGTCAGAACCTTGCTGCTATATTTCACACTGCTACACTTTCCTGGGCTGGTTC 60

|||||  
214 ATCAGGTCAGAACCTTGCTGCTATATTTCACACTGCTACACTTTCCTGGGCTGGTTC 273

QY 61 CCAGAAAGATTCTGAGCAATCACTATTTGCGATGCGCACTATACAAACAGCATCCGGTG 120



QY	2281	ACCCCAACGCTCGACGAGAAATGGGAAGCAGCCGCGGAGAGGCCGAGTGTGAGAGAGAAC	2349
DB	2494	AACCCAAAGCTTCAGTGAAGAAGGGGAACGCGGGGACAGCGGAGTGGGAGTAGACAAC	2563
QY	2341	CAGAAGCTCATCAATGCCTGSCAATAAGACATGGCCCCCATGGCTCCGCTGTGATTCCC	2400
DB	2554	CAGAAGCTCATCAATGCCTGSCAATAAGACATGGCCCCCATGGCTCCGCTGTGATTCCC	2613
QY	2401	ACGGAGCTCGCCTGGGGCTCTCCCGACGCACATCCCGAGGGTGTTGCTCTGGCCATC	2460
DB	2614	ACGGAGCTCGCCTGGGGCTCTCCCGACGCACATCCCGAGGGTGTTGCTCTGGCCATC	2673
QY	2461	ACGCACAGGGCTACCAGACATAGTAGCTACCTGGACAGCCCAAAATGAGCAGGTGGCCAG	2520
DB	2674	ACGCACAGGGCTACCAGACATAGTAGCTACCTGGACAGCCCAAAATGAGCAGGTGGCCAG	2733
QY	2521	ATGGCGCTGGAGGACACAGCGCGCCACTGGAGTAAGAACCATCAAGGAACATCTACGC	2580
DB	2734	ATGGCGCTGGAGGACACAGCGCGCCACTGGAGTAAGAACCATCAAGGAACATCTACGC	2793
QY	2581	ASGACAGTCCCAACCATGGGTGAACCTGTGGAGANCTGGACAGCTGGCCCCCAA	2640
DB	2794	ASGACAGTCCCAACCATGGGTGAACCTGTGGAGANCTGGACAGCTGGCCCCCAA	2853
QY	2641	GTTCCACAGCGGAGGCTCTCCCTGGGTCCCGCGGAGCGTCCCTGTCTCAGACCGGTCTA	2700
DB	2854	GTTCCACAGCGGAGGCTCTCCCTGGGTCCCGCGGAGCGTCCCTGTCTCAGACCGGTCTA	2913
QY	2701	AGCAAGCGGCTGAAATGCAACACTCCTCTCTCTAGCGGGTTGACTATAGAGGAGCTAC	2760
DB	2914	AGCAAGCGGCTGAAATGCAACACTCCTCTCTCTAGCGGGTTGACTATAGAGGAGCTAC	2973
QY	2761	CCACAGAACTCGCTACGAGAGGCCACAGGCCACG - ACTCTCAAAGAAACCAACTAA	2819
DB	2974	CCACAGAACTCGCTACGAGAGGCCACAGGCCACG - ACTCTCAAAGAAACCAACTAA	3033
QY	2820	CTCCTCCAATCTCTCTC - AACTCTCAGAAACACAGAGCTTTGACGGGAGACACCCGC	2878
DB	3034	CCCGAGNATCANCTCTGACTTCAAGGAGCACAGAGCTTTGACGGGAGACACCCGC	3093
QY	2879	CGCCGCCCGCAGAGGGTGSAATCCATTCAGGTGCACAGCTCCCAAGCATCTGSGCCAGG	2938
DB	3094	CGCCGCCCGCAGAGGGTGSAATCCATTCAGGTGCACAGCTCCCAAGCATCTGSGCCAGG	3153
QY	2939	CCGTGACTCTCTGAGGAGACCCAGCGCTCAAGCGCTACAACTCATGACAGGTCTGGGCG	2998
DB	3154	CCGTGACTCTCTGAGGAGACCCAGCGCTCAAGCGCTACAACTCATGACAGGTCTGGGCG	3213
QY	2999	TGAAGGTACGCGCTCGCTAAGCGCGAGCTACCCCGCAAGCACTCTTGTGCTCCCTTT	3058
DB	3214	TGAGGCTACGCGCTCGCTAAGCGCGAGCTACCCCGCAAGCACTCTTGTGCTCCCTTT	3273
QY	3059	CCACATCCATGAAGCCCAATGATGCGTGTACATAA	3093
DB	3274	CCACATCCATGAAGCCCAATGATGCGTGTACATAA	3308

## RESULT 2

RESULTS 2  
US-09-957-187-3

US-05-937-107-3 ; Sequence 3, Application US/09951  
; Publication No. US20030054514A1

; GENERAL INFORMATION:

APPLICANT: Shimkets, Richard A.

APPLICANT: LaRochelle, William

THE INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

FILE REFERENCE: 15908 340 CIF  
; CURRENT APPLICATION NUMBER: US/09/957,187

; CURRENT APPLICATION NUMBER: US/09/337,187  
 ; CURRENT FILING DATE: 2000-09-19

: CURRENT FILING DATE: 2000 03 13  
 : PRIOR APPLICATION NUMBER: 60/123 567

PRIOR FILING DATE: 1999-03-09

PREVIOUS FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/520,781

; PRIOR FILING DATE: 2000-03-23

CONFIDENTIAL

```

; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (214)..(3030);
; NAME/KEY: misc_feature
; LOCATION: (3047)
; OTHER INFORMATION: an n may be any one of a or t or g or c
US-09-457-167-3

Query Match: 98.2%; Score 3038; DB 11; Length 3498;
Best Local Similarity 99.3%; Pred No. 0;
Matches 3072; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

QY 1 ATGAGGTCAGAGCCCTTGCCTATATTTCCACACTGCTACACTTTCGTGGCGTGTTC 60
DB |||||||
DB 214 ATGAGGTCAGAGCCCTTGCCTATATTTCCACACTGCTACACTTTCGTGGCGTGTTC 273
QY 61 CCAGAGATCTCAGGCCAATCAGTATTTCCCATGGCAACTATACAAACAGTATCCGGIG 120
DB |||||||
DB 274 CCAGAGATCTCAGGCCAATCAGTATTTCCCATGGCAACTATACAAACAGTATCCGGIG 333
QY 121 TTITGGGCCAACAGCCAGGACGGAACACACACAGAGGCACAGGCTGGACATCCAGATG 180
DB |||||||
DB 334 TTITGGGCCAACAGCCAGGACGGAACACACACAGAGGCACAGGCTGGACATCCAGATG 393
QY 181 ATTATGATCATGAACGGAGCCCTCTCATTTGCTAGGACCAATATTTTACTGTGTGAT 240
DB |||||||
DB 394 ATTATGATCATGAACGGAGCCCTCTCATTTGCTAGGACCAATATTTTACTGTGTGAT 453
QY 241 ATAGACACATCACACACGGAGAAATTTATTGTAGCAAAAACATGCATGGAATCTAGA 300
DB |||||||
DB 454 ATAGACACATCACACACGGAGAAATTTATTGTAGCAAAAACATGCATGGAATCTAGA 513
QY 301 CAGGCCGATGAGACATCAGAGATGAAGGGAAGAAATAGGATGAGTCCACAGCTTT 360
DB |||||||
DB 514 CAGGCCGATGAGACATCAGAGATGAAGGGAAGAAATAGGATGAGTCCACAGCTTT 573
QY 361 ATTAAGATCTCTAAAGAAACAGTATGCATTTGTTCTGTGGAACATATGCGCTTC 420
DB |||||||
DB 574 ATTAAGATCTCTAAAGAAACAGTATGCATTTGTTCTGTGGAACATATGCGCTTC 633
QY 421 AACGCTTCTGCAGAAACTATAGATGGATACATTTGGAACCATTCGGGATGATTCAGC 480
DB |||||||
DB 634 AACGCTTCTGCAGAAACTATAGATGGATACATTTGGAACCATTCGGGATGATTCAGC 693
QY 481 GGAATGCCAGATGCCCATATGTCGCAACATGCCAACGTTGCACATGTTTGACAGATGGA 540
DB |||||||
DB 694 GGAATGCCAGATGCCCATATGTTGCCAACAATGCCAACGTTGTCACATGTTTGACAGATGGA 753
QY 541 AAACATACTCAGCCACAGTACATGACTTTCCTTGCCATTGACGCAGTCAATTACCGGAGT 600
DB |||||||
DB 754 AAACATACTCAGCCACAGTACATGACTTTCCTTGCCATTGACGCAGTCAATTACCGGAGT 813
QY 601 CTTGGAGAAAGCCCTACGCTCGGACCGCTCAAGCAGCATTCAAATGGTTGAAGAAGACCA 660
DB |||||||
DB 814 CTTGGAGAAAGCCCTACGCTCGGACCGCTCAAGCAGCATTCAAATGGTTGAAGAAGACCA 873
QY 661 TACTTTGTTCAAGCGGTGGATTACGGAGATATATCTACTTCTTTTCACGGGAATAGCA 720
DB |||||||
DB 874 TACTTTGTTCAAGCGGTGGATTACGGAGATATATCTACTTCTTTTCACGGGAATAGCA 933
QY 721 GTGAGTATTAACACCATGGAAAGGTAGTTTTCGCAAGAGTGGCTCAGTTTGTGAAGAT 780

```

DB 934 GTGAG-ATAACACCATGGGAAGGTAGTTTCCAGAGTGGCTCAGGTTTSTAAGAT 943  
QY  
DB 781 GATATGGAGATCTCAAGAGTCTGGAGAAAGTGGACCTGGTTCGTGAAGGCGGC 840  
DB 994 GATATGGAGATCTCAAGAGTCTGGAGAAAGTGGACCTGGTTCGTGAAGGCGGC 1053  
QY 841 TGAAGTGGTCAAGTTCCTGGAGTCTCAATTTTATTTCAACATCTCCAGGCACTTGA 900  
DB 1054 TGAAGTGGTCAAGTTCCTGGAGTCTCAATTTTATTTCAACATCTCCAGGCACTTGA 1113  
QY 901 GATGATGATGATCAAGAGTGGTGAATTTGCTGGACAGTTCCTACACCTATTAAG 960  
DB 1114 GATGATGATGATCAAGAGTGGTGAATTTGCTGGACAGTTCCTACACCTATTAAG 1173  
QY 961 AGCATGCTGGTCTGGAGTCTGGCTATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 1174 AGCATGCTGGTCTGGAGTCTGGCTATGATGATGATGATGATGATGATGATGATGAT 1233  
QY 1021 GGGAGATTCAGAGAAAGATCTCGGATTCAGCTGGAGACACCTCTGATGAGGA 1080  
DB 1234 GGGAGATTCAGAGAAAGATCTCGGATTCAGCTGGAGACACCTCTGATGAGGA 1093  
QY 1081 GTTCTAAGCTCGAGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1140  
DB 1294 GTTCTAAGCTCGAGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1353  
QY 1341 TCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1200  
DB 1354 TCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1413  
QY 1201 GAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1260  
DB 1414 GAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1473  
QY 1261 ACCAATTCAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 1474 ACCAATTCAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1533  
QY 1321 GGATCAGAGAGGGAATCATCTGAAGTCTTGGGAGATAGGAATAGTGGTTCCTA 1380  
DB 1534 GGATCAGAGAGGGAATCATCTGAAGTCTTGGGAGATAGGAATAGTGGTTCCTA 1593  
QY 1381 AATCAGAGCTTTTCCTGGAGAGATGATGATGATGATGATGATGATGATGATGATG 1440  
DB 1594 AATCAGAGCTTTTCCTGGAGAGATGATGATGATGATGATGATGATGATGATGATG 1653  
QY 1441 GGAGTCCGAAGCAAAAGGATCATGGGATGAGTGGAGACCAAGCAAGCTCTCTGAT 1500  
DB 1654 GGAGTCCGAAGCAAAAGGATCATGGGATGAGTGGAGACCAAGCAAGCTCTCTGAT 1713  
QY 1501 GTTGGTCTCTAGCTGTGTGATTAAGGTTCCGCTTGGCGGTGTGAAGACATGGGAAG 1560  
DB 1714 GTTGGTCTCTAGCTGTGTGATTAAGGTTCCGCTTGGCGGTGTGAAGACATGGGAAG 1773  
QY 1561 TGTAAAAAAGCTGTATGTGCTCCAGAGACCCATATGTGGATGGATTAAGGAGGTGGT 1620  
DB 1774 TGTAAAAAAGCTGTATGTGCTCCAGAGACCCATATGTGGATGGATTAAGGAGGTGGT 1833  
QY 1621 CCTCGAGAGCTTTATCAGCCAGACAGACCTGACCTTGAAGAGAGATAGAGGCTGGC 1680  
DB 1834 CCTCGAGAGCTTTATCAGCCAGACAGACCTGACCTTGAAGAGAGATAGAGGCTGGC 1893  
QY 1681 AATCAGATGCTGGGAGTGTGACAAATTTGCTGGACGTGAATGGGATTCAGT 1740  
DB 1894 AATCAGATGCTGGGAGTGTGACAAATTTGCTGGACGTGAATGGGATTCAGT 1953  
QY 1741 TCCCTCTTCCCGACCAACACATCAGATTCAGAGCTTAAGAGGCTATGATGATGATG 1800  
DB 1954 TCCCTCTTCCCGACCAACACATCAGATTCAGAGCTTAAGAGGCTATGATGATGATG 2013  
QY 1801 GGAGAAATGCTGGAGTGGAGATCTGCTTGAATCAGCTGACAGACAGACCTTTGGG 1860  
DB

DB 2014 GGAGAAATGCTGGAGTGGAGATCTGCTTGAATCAGCTGACAGACAGACCTTTGGG 2073  
QY 1861 GCACTGCTTCCCATATATCATCAAGACAAGAGGAGTGAITGGGAAAGTACCTCAAA 1920  
DB 2074 GCACTGCTTCCCATATAACACAGACAAGAGGAGTGAITGGGAAAGTACCTCAAA 2133  
QY 1921 GGCACAGACAGCTGGTTCGGGTCAGCTCTTTGGGCACTGACATCTGCTGGCTTCGTC 1980  
DB 2134 GGCACAGACAGCTGGTTCGGGTCAGCTCTTTGGGCACTGACATCTGCTGGCTTCGTC 2193  
QY 1981 ATGGGCGGCTTCTCGGCTCAGCTCAGCTCTGCTGCTGTGATCATGGGCAAGAG 2040  
DB 2194 ATGGGCGGCTTCTCGGCTCAGCTCAGCTCTGCTGCTGTGATCATGGGCAAGAG 2253  
QY 2041 GTGGCTGTGGTGGAGCAAGAGAGAGGAGTCTCAGCTCTCGGCGGGGCTCCATGAGC 2100  
DB 2254 GTGGCTGTGGTGGAGCAAGAGAGAGGAGTCTCAGCTCTCGGCGGGGCTCCATGAGC 2313  
QY 2101 AGCTCAGCAAGCTCAGGCGCTTTTGGGCACTCAATCCAAAGACCAAGCGGAG 2160  
DB 2314 AGCTCAGCAAGCTCAGGCGCTTTTGGGCACTCAATCCAAAGACCAAGCGGAG 2373  
QY 2161 GGCATGCTCAGGCGCTCAATGCAAGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
DB 2374 GGCATGCTCAGGCGCTCAATGCAAGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2433  
QY 2221 ATGCTGATTAAGCAGACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2280  
DB 2434 ATGCTGATTAAGCAGACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2493  
QY 2281 ACCCAAGCTTGCAGAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340  
DB 2494 ACCCAAGCTTGCAGAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2553  
QY 2341 CAGAGCTCAATGCTGCTCAGAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
DB 2554 CAGAGCTCAATGCTGCTCAGAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2613  
QY 2401 AGGAGCTTGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
DB 2614 AGGAGCTTGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2673  
QY 2461 AGGAGAGGCTTACCAGATGAGTACGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 2520  
DB 2674 AGGAGAGGCTTACCAGATGAGTACGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 2733  
QY 2521 ATGGGCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580  
DB 2734 ATGGGCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2793  
QY 2581 AGCAAGAGTCCCAAGCTTGGGCTGAGCTTGTGGAGAGCTGAGAGCTGAGAGCTGAGAG 2640  
DB 2794 AGCAAGAGTCCCAAGCTTGGGCTGAGCTTGTGGAGAGCTGAGAGCTGAGAGCTGAGAG 2853  
QY 2641 GTTCCAGAGGAGGAGGCTTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
DB 2854 GTTCCAGAGGAGGAGGCTTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2913  
QY 2701 AGCAAGGAGGCTTGGAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
DB 2914 AGCAAGGAGGCTTGGAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2973  
QY 2761 CCCAGAACTGCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2819  
DB 2974 CCCAGAACTGCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3033  
QY 2820 CTCCTCCAAATTCCTCTC-ACCTCTCCAAAGAGAGCTTGGAGGAGAGAGAGAGAGAG 2878  
DB 3034 CCCGAGAAATTCCTCTGATTTCAAAGGAGGAGAGCTTGGAGGAGAGAGAGAGAGAGAG 3093  
QY 2879 GCGCGGCGCGGAGAGGAGTGGAGTCCATCAGAGTGGAGAGCTCCAGAGCTTGGCGAGG 2938  
DB 3094 GCGCGGCGCGGAGAGGAGTGGAGTCCATCAGAGTGGAGAGCTCCAGAGCTCCAGAGCT 3153

Qy	2939	CCCTGACTGTCTGTAGAGACACCCAGAGCTTAAAGCCCTACACCTCATGTAGCAAGGTGAGGAG	2949
Db	3154	CCCTGACTGTCTGTAGAGACACCCAGAGCTTAAAGCCCTACACCTCATGTAGCAAGGTGAGGAG	3214
Qy	2999	TGAAGAGTAGCGGCTGCTTAAGACCGGAGCTAGCCGCCCAAGCATGCTTTGCTCCGCTTT	3056
Db	3214	TGAGGGTACGGCTGCTTAAAGCCCGGAGCTAGCCGCCCAAGCATGCTTTGCTCCGCTTT	3273
Qy	3059	CCACATCCATAGAGGCCAAATGATGGCTGTACATAA	3093
Db	3274	CCACATCCATAGAGGCCAAATGATGGCTGTACATAA	3308

RESULT 3

US-09-957-187-84

; Sequence 84, Application US/09957187

; Publication No. US20030054514A;

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard A.

; APPLICANT: LeRoche, William

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY

; FILE REFERENCE: 15966-540 CIP

; CURRENT APPLICATION NUMBER: US/09/957,187

; CURRENT FILING DATE: 2000-09-19

; PRIOR APPLICATION NUMBER: 60/123,667

; PRIOR FILING DATE: 1999-03-09

; PRIOR APPLICATION NUMBER: 09/520,781

; PRIOR FILING DATE: 2000-03-03

; PRIOR APPLICATION NUMBER: 60/234,082

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: 63/233,798

; PRIOR FILING DATE: 2000-09-19

; PRIOR APPLICATION NUMBER: 60/174,485

; PRIOR FILING DATE: 2000-01-04

; NUMBER OF SEQ ID NOS: 85

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 84

; LENGTH: 4250

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (250)..(3390)

US-09-957-187-84

Query Match 98.0%; Score 3030.4; DB 11; Length 4250;

Best Local Similarity 98.3%; Pred. No. 0;

Matches 3092; Conservative 0; Mismatches 1; Indels 51; Gaps 17

Qy	1	ATCAGGTCCACAAGCGCTTGGCTGCTATATTTCACACTGCTACACTTGGTGGGGCGTGTTC	60
Db	250	ATCAGGTCCACAAGCGCTTGGCTGCTATATTTCACACTGCTACACTTGGTGGGGCGTGTTC	309
Qy	61	CCACAAGATTCTGAGCCGATCATGTAATTTGGATGCGGCAATATACAAACAGATTCGGGTG	120
Db	310	CCAGAAGATTCTGAGCCGATCATGTAATTTGGATGCGGCAATATACAAACAGATTCGGGTG	369
Qy	121	TTTGTGGGTCACAAGCCAGCGCCGACACACACAGAGGCGACGGCTGGACATCCAGATG	180
Db	370	TTTGTGGGTCACAAGCCAGCGCCGACACACACAGAGGCGACGGCTGGACATCCAGATG	420
Qy	181	ATTATGATCATGTAGCCGAGCGGCTGATGCTGCTAGCGACCAATATTTAATCTTGAT	240
Db	430	ATTATGATCATGTAGCCGAGCGGCTGATGCTGCTAGCGACCAATATTTAATCTTGAT	489
Qy	241	ATGACACATCCACACGCGGAGGAATTTATTTGTAGCAAAAATGTGCATGGAAATCTAGA	500
Db	490	ATGACACATCCACACGCGGAGGAATTTATTTGTAGCAAAAATGTGCATGGAAATCTAGA	549
Qy	301	CAGGCCGATCTAGACACATGCACAATGCARGGGAAACATAGGATGAGTGCCACAACCTT	360
Db	550	CAGGCCGATCTAGACACATGCACAATGCARGGGAAACATAGGATGAGTGCCACAACCTT	609

QY	361	ATTAAAGTCTTCTTAAGAAAAACGATGATGCAATGTGTTCTGTGSAACAAATGCGTTTC	420
DB	610	ATTAAAGTCTTCTTAAGAAAAACGATGATGCAATGTGTTCTGTGSAACAAATGCGTTTC	669
QY	421	ANCCGTCTCTGAGAAACATACATGGATGATACATTGGAAACCAATGCGGGGATGAATCAGC	480
DB	670	ANCCGTCTCTGAGAAACATACATGGATGATACATTGGAAACCAATGCGGGGATGAATCAGC	729
QY	481	GGAAATGCCCCAGAAGCCCATATGAAGGCAAAACATGCCCCAATGCTACTGTTGCGAGATGGA	540
DB	730	GGAAATGCCCCAGATGCCCATATGATGCAAAACATGCCCAACGTTGCACTGTTGCGAGATGGA	789
QY	541	AAACTATACTACGCCACACGATGACTGACTTCCTTGGCAATTCAGCCAGCTCATTTTACCGGAGT	600
DB	790	AAACTATACTACGCCACACGATGACTTCCTTGGCAATTCAGCCAGCTCATTTTACCGGAGT	849
QY	601	CTTGGAAAAAGCCGTACCCCTGGGCACCGCTCAAGCACGATTCAAAATGGTTGAAAGAACCA	660
DB	850	CTTGGAAAAAGCCGTACCCCTGGGCACCGCTCAAGCACGATTCAAAATGGTTGAAAGAACCA	909
QY	661	TACTTTGTTCAAGGCGGNGATTAACGGAGATTATATCTACTTCTCTTCACGGAAATAGCA	720
DB	910	TACTTTGTTCAAGGCGGNGATTAACGGAGATTATATCTACTTCTCTTCACGGAAATAGCA	969
QY	721	GTGGASTATAACACCATGSGGAAAGGTAGTTTTGCCAAGAGTGGCTCAGSTTTGTGAAGAT	780
DB	970	GTGGAGTATAACACCATGSGGAAAGGTAGTTTTGCCAAGAGTGGCTCAGSTTTGTGAAGAT	1029
QY	781	GATATGGGAGATCTCAAGAGTCTCTGGAGAACAGTGGAGCTGCTTCCTTGAAGGCGGC	840
DB	1030	GATATGGGAGATCTCAAGAGTCTCTGGAGAACAGTGGAGCTGCTTCCTTGAAGGCGGC	1089
QY	841	TTGAACCTGCTCAGTTCCTGGAGACTCTCATTTTTTATTTCAACATCTCCACGGCAGTTTACA	900
DB	1090	TTGAACCTGCTCAGTTCCTGGAGACTCTCATTTTTTATTTCAACATCTCCACGGCAGTTTACA	1149
QY	901	CATGTGATTCGTATCAACGGGSGTGATGTTGTCTGAGCAACAGTTTCTACACCTTATTAAC	960
DB	1150	CATGTGATTCGTATCAACGGGSGTGATGTTGTCTGAGCAACAGTTTCTACACCTTATTAAC	1209
QY	961	AGCATCCTGGGTATCCAGTCTGTGCCATTATGCATGCTTGACATTTGCCAGTGTGTTTACT	1020
DB	1210	AGCATCCTGGGTATCCAGTCTGTGCCATTATGCATGCTTGACATTTGCCAGTGTGTTTACT	1269
QY	1021	CGAGATTTCAAGAACAGAGACTCTCCGATTCACCTTGGACACCAAGTTCCTGATGAACGA	1080
DB	1270	CGAGATTTCAAGAACAGAGACTCTCCGATTCACCTTGGACACCAAGTTCCTGATGAACGA	1329
QY	1081	GTTCCTAAGCCCCAGGCCAGGTTCGTGTGCTGGCTCATCTCTCTAGAAAAGATATGCAACC	1140
DB	1330	GTTCCTAAGCCCCAGGCCAGGTTCGTGTGCTGGCTCATCTCTCTAGAAAAGATATGCAACC	1389
QY	1141	TCGNATGAGTTCCCTGATGATACCGCTCAACTTCATCAAGACCGCACCGCTCTGCGATGAG	1200
DB	1390	TCGNATGAGTTCCCTGATGATACCGCTCAACTTCATCAAGACCGCACCGCTCTGCGATGAG	1449
QY	1201	GCAGTGCCTTCCATCTTCAACAGGCCATGGTTCCTGAGAACAAATGGTCAGATACCGGCTT	1260
DB	1450	GCAGTGCCTTCCATCTTCAACAGGCCATGGTTCCTGAGAACAAATGGTCAGATACCGGCTT	1509
QY	1261	ACCAAAATTCAGTGGACACAGCTGCTGGGCCATATTCAGAAATCACACTGTGGGTTTTCTG	1320
DB	1510	ACCAAAATTCAGTGGACACAGCTGCTGGGCCATATTCAGAAATCACACTGTGGGTTTTCTG	1569
QY	1321	GGATCAGAGAGGGAATCACTCTTGAAGTTTTTGGCCAGATAGAAATAGTGGTTTTCTA	1380
DB	1570	GGATCAGAGAGGGAATCACTCTTGAAGTTTTTGGCCAGATAGAAATAGTGGTTTTCTA	1629
QY	1381	AATGACAGCCTTTTCCTGGAGGAGATGAGTGTATTACAACTCTGAAAAATGCAGCTATGAT	1440
DB	1630	AATGACAGCCTTTTCCTGGAGGAGATGAGTGTATTACAACTCTGAAAAATGCAGCTATGAT	1689





Query Match	47.24	Score 2696	DP 11	Length 3333	
Best Local Similarity	93.94	Prod. No. 0			
Matches 2907	Conservative	C	Mismatches 21	Indels 167	Gaps
QY	1	ATGAGCTCAGAGGCTTGTGCTATATATTCACATCTGTTACACTTTCCTGGGCTGCTTTC	60		
DB	214	ATGAGCTCAGAGGCTTGTGCTATATTCACACGCTACACTTTCCTGGGCTGCTTTC	273		
QY	61	CCAGAAATTTCTGAGCCAGACATATTTTCGATGTGCAACTACAAAAACAGTATCCGCTG	120		
DB	274	CCAGAAATTTCTGAGCTAATCAGATTTTCGATGTGCAACTACAAAAACAGTATCCGCTG	333		
QY	121	TTTGTGGGCCACAGCCAGGACGAGACACACACAGAGGACAGGCTGGACATCCAGATG	180		
DB	334	TTTGTGGGCCACAGCCAGGACGAGACACACACAGAGGACAGGCTGGACATCCAGATG	393		
QY	181	ATTATGATCATGAAGCAAGCCCTCTACATTGCTGTAGGACCAATTATATCTGTGAT	240		
DB	394	ATTATGATCATGAAGCAAGCCCTCTACATTGCTGTAGGACCAATTATATCTGTGAT	453		
QY	241	ATAGACACATCACACGCGAAGAAATTTATTTAGCAAAAACACTGACATGGAAATCTAGA	300		
DB	454	ATAGACACATCACACGCGAAGAAATTTATTTAGCAAAAACACTGACATGGAAATCTAGA	513		
QY	301	CAGCCGATGTAGACACATCGAGATGAAGGGAANAACATAGAGTACGTGCGCAACATTT	360		
DB	514	CAGCCGATGTAGACACATCGAGATGAAGGGAANAACATAGAGTACGTGCGCAACATTT	573		
QY	361	ATTTAAAGTTCTCTAAAGAAAAACGATGATGATTTTGTCTGTGGAACCTAATSCCTTC	420		
DB	574	ATTTAAAGTTCTCTAAAGAAAAACGATGATGATTTTGTCTGTGGAACCTAATSCCTTC	633		
QY	421	TACCTTCTGCAAAACTATAGATGGGATAGATTTGGAAACCATTCGGGGATGAATTCAGC	480		
DB	634	TACCTTCTGCAAAACTATAGATGGGATAGATTTGGAAACCATTCGGGGATGAATTCAGC	593		
QY	481	GAATGGCCAGATGCCCATATATGTCGCAACCATGCCAAGTTTGCATTTTTCAGATGGA	540		
DB	694	GAATGGCCAGATGCCCATATATGTCGCAACCATGCCAAGTTTGCATTTTTCAGATGGA	753		
QY	541	AAACTATCTACGACAGTACTGACTTCTCTGCACTTACGCGACTCATTTACCGAGTT	600		
DB	754	AAACTATCTACGACAGTACTGACTTCTCTGCACTTACGCGACTCATTTACCGAGTT	813		
QY	601	CTTGGAANAAGCTTACTCTGCGAGAGCTGCAAGCAGATTCAAAATGTTTGNAAAGACCA	660		
DB	814	CTTGGAANAAGCTTACTCTGCGAGAGCTGCAAGCAGATTCAAAATGTTTGNAAAGACCA	873		
QY	661	TACTTTGTTCAAGCTTGGATACGAGATATATCTACTTCTCTTTCAGGGAATATAGCA	720		
DB	874	TACTTTGTTCAAGCTTGGATACGAGATATATCTACTTCTCTTTCAGGGAATATAGCA	933		
QY	721	GTGGAGTATACACCATGCGAAAGTAGTTTTCGCAAGAGTCTGTCAGTTTATAGAAT	780		
DB	934	GTGGAGTATACACCATGCGAAAGTAGTTTTCGCAAGAGTCTGTCAGTTTATAGAAT	993		
QY	781	GATATGGGAGGATCTCAAGAGTCTGTGGAGAAACAGAGAGCTCTTCTCTCAAGCGGCGC	840		
DB	994	GATATGGGAGGATCTCAAGAGTCTGTGGAGAAACAGAGAGCTCTTCTCTCAAGCGGCGC	1053		
QY	841	TTTAACTCTCAGTTTCCTGGAGACTCTCATTTTATTTTCAACATTCTTCAGGCGACTTACA	900		
DB	1054	TTTAACTCTCAGTTTCCTGGAGACTCTCATTTTATTTTCAACATTCTTCAGGCGACTTACA	1113		
QY	901	GATGTAATTCATCAAGGGGCGATGATGTTCTGCTGTGGCAAGCTTTTCTACACCTTAATAC	960		
DB	1114	GATGTAATTCATCAAGGGGCGATGATGTTCTGCTGTGGCAAGCTTTTCTACACCTTAATAC	1173		
QY	961	ASCATCCTGGGCTGCGAGTCTGTGCGCTATGACATGCTTGACATTCGCCAGTCTTTTACT	1020		
DB	1174	ASCATCCTGGGCTGCGAGTCTGTGCGCTATGACATGCTTGACATTCGCCAGTCTTTTACT	1233		

Qy	1621	GGAGATTC	TAAGGAACAGAAAGTCTCTCCTGATTCACACCCGACACAGATCTCTGATGACGA	1081
Dz	1624	GGGAGATTC	TAAGGACAGAAAGTCTCTCTGATTCGACCTGGACACGATTCGCTGATGACGA	1293
Qy	1681	GTTCCTTAAGCC	CAAGGSCAGGTGCTGTGCTGGCTCATCTCTCTTACAAAGATATGCAACC	1140
Dz	1294	GTTCCTTAAGCC	CAAGGSCAGGTGCTGTGCTGGCTCATCTCTCTTACAAAGATATGCAACC	1353
Qy	1341	TCCAAATGAGTTC	CCCTATATGATACCCCTGAATTTCTATCAAGAGCAGCCGCTCATCTGATGAG	1300
Dz	1354	TCCAAATGAGTTC	CCCTATATGATACCCCTGAATTTCTATCAAGAGCAGCCGCTCATCTGATGAG	1413
Qy	1201	GCAGTGGCCCT	ATCTCTTACAGAGCCCATGGTTCCTGAGACAAATGGTCAGATACCGCTT	1260
Dz	1414	GCAGTGGCCCT	ATCTCTTACAGAGCCCATGGTTCCTGAGACAAATGGTCAGATACCGCTT	1473
Qy	1261	ACCAAAATTCG	AGTGCACACAGCTGCTGGGCCATATCAGAAATCACACTGTGGTTTTCCTG	1320
Dz	1474	ACCAAAATTCG	AGTGCACACAGCTGCTGGGCCATATCAGAAATCACACTGTGGTTTTCCTG	1533
Qy	1321	GGATCAGAGGGA	ATCATCTTTGAAGTTTTTGGCCAGAAATAGGAATATGTTGTTTCTA	1380
Dz	1534	GGATCAGAGGGA	ATCATCTTTGAAGTTTTTGGCCAGAAATAGGAATATGTTGTTTCTA	1593
Qy	1381	AATGCACAGCCTT	TCCTGGAGGAGATGAGTGTTTACAACTCTGAAAATGCAGCTATGAT	1440
Dz	1594	AATGCACAGCCTT	TCCTGGAGGAGATGAGTGTTTACAACTCTGAAAATGCAGCTATGAT	1653
Qy	1441	GGAGTGAAGACAA	AAAGGATCATGGGCATGCAGCTGGACAGACAGCAGCTCTCTCTGAT	1500
Dz	1654	GGAGTGAAGACAA	AAAGGATCATGGGCATGCAGCTGGACAGACAGCAGCTCTCTCTGAT	1713
Qy	1501	GTTCGGTCTCT	ACTCTGTGTGATAAAGGTCCCCCTTGGCCGGTGTGAACACATGCGAAG	1560
Dz	1714	GTTCGGTCTCT	ACTCTGTGTGATAAAGGTCCCCCTTGGCCGGTGTGAACACATGCGAAG	1773
Qy	1561	TGTAAAAAACCT	GTATTTGCTTCAGAGACCCATATTTGTGTGATGGATTAAGGAAGGTGGT	1620
Dz	1774	TGTAAAAAACCT	GTATTTGCTTCAGAGACCCATATTTGTGTGATGGATTAAGGAAGGTGGT	1833
Qy	1621	GCCTCGACCCAT	TTATCACCCACACAGCAGACTGACTTTTGACGAGGACATAGACGCTGAC	1680
Dz	1834	GCCTCGACCCAT	TTATCACCCACACAGCAGACTGACTTTTGACGAGGACATAGACGCTGAC	1893
Qy	1681	AATACAGATGCT	CGGAGACTGCACAAATCCCTTTTGCCACTGGAATGGCATTCCAGT	1740
Dz	1894	AATACAGATGCT	CGGAGACTGCACAAATCCCTTTTGCCACTGGAATGGCATTCCAGT	1941
Qy	1741	TCCCTCTTGCC	ACACACATCAGATTCGACGGCTCAAGAGGCGTATGAGTCTAGG	1800
Dz	1942	TCCCTCTTGCC	ACACACATCAGATTCGACGGCTCAAGAGGCGTATGAGTCTAGG	1941
Qy	1801	GGAGGAAATGCT	GTGACTGGAAGCATCTGCTTTGACTCACCCTGACACACAGACCCCTTGGG	1860
Dz	1942	GGAGGAAATGCT	GTGACTGGAAGCATCTGCTTTGACTCACCCTGACACACAGACCCCTTGGG	1941
Qy	1861	GCAGTGTCTT	CCCATATCATCCACAGACAGAGGAGTGGAGTGGTTCGGGAAAGTTACCTCAA	1920
Dz	1942	GCAGTGTCTT	CCCATATCATCCACAGACAGAGGAGTGGAGTGGTTCGGGAAAGTTACCTCAA	1968
Qy	1921	GGCCACGACGAG	CTGGTTCGGCTCAGCCCTCTTGCCCATTTGACGTCATCCCTGGCTTCGCT	1980
Dz	1969	GGCCACGACGAG	CTGGTTCGGCTCAGCCCTCTTGCCCATTTGACGTCATCCCTGGCTTCGCT	2028
Qy	1981	ATGGGGGCCG	TCTTCTCGGATACACCGTCTACTGCGCTCTGTGTATCATCGGCGCAAGAC	2040
Dz	2029	ATGGGGGCCG	TCTTCTCGGATACACCGTCTACTGCGCTCTGTGTATCATCGGCGCAAGAC	2088
Qy	2041	GTGGCTGTGG	TGCAGCGCAAGGAGAGGAGCTACCCACTCGCGCCGGGCTCCATGAGC	2100
Dz	2089	GTGGCTGTGG	TGCAGCGCAAGGAGAGGAGCTACCCACTCGCGCCGGGCTCCATGAGC	2148
Qy	2101	AGCCTCACCA	AGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAG	2160

Db 2149 AGCGCACCAAGCTCAGCGGCTCTCTGTGGGAGACCAATCCAAAGACCCAAAGCCGAG 2208  
QY 2161 GCCATGCTCAGCGGACCTCATCAGCAAGCGGAAGTCCCACTCCCGGCAACAGCGGCAAG 2220  
Db 2209 CCCATCTCAGCGGACCTCATCAGCAAGCGGAAGTCCCACTCCCGGCAACAGCGGCAAG 2266  
QY 2221 ATGCTCATTAAGCAGACAGCAGCAGCAGCTGAGCTGACGGCCCTCCCAACCCAGAGTCA 2269  
Db 2269 ATGCTCATTAAGCAGACAGCAGCAGCAGCTGAGCTGACGGCCCTCCCAACCCAGAGTCA 2328  
QY 2281 ACCCAACGCTGCAGCAGAGCGGAAGCCAGCGCGGCAAGCGGAGTGGGAGAGGAC 2340  
Db 2329 ACCCAACGCTGCAGCAGAGCGGAAGCCAGCGCGGCAAGCGGAGTGGGAGAGGAC 2388  
QY 2341 CAGAACCTCATCAATGCTGTCAAGAGGATGCGCCGCAATGGCTGCTGCTGATTCG 2400  
Db 2399 CAGAACCTCATCAATGCTGTCAAGAGGATGCGCCGCAATGGCTGCTGCTGATTCG 2449  
QY 2401 ACGGACCTGCTGCTGCGGCGCTCCCGCAGCAGCAGATCCGAGCGTGTGTGCTGCGCATC 2460  
Db 2449 ACGGACCTGCTGCTGCGGCGCTCCCGCAGCAGCAGATCCGAGCGTGTGTGCTGCGCATC 2508  
QY 2461 ACGGAGAGGCTTACAGCAAGAGTGTGCTGACAGAGCCCAAAATGAGCGAGGTGGCCAG 2520  
Db 2509 ACGGAGAGGCTTACAGCAAGAGTGTGCTGACAGAGCCCAAAATGAGCGAGGTGGCCAG 2568  
QY 2521 ATGGCGCTGGAGAGAGCGGCGGCGGCAAGCTGAGTATAGAGAGATCAGGAGATCTCAG 2580  
Db 2569 ATGGCGCTGGAGAGAGCGGCGGCGGCAAGCTGAGTATAGAGAGATCAGGAGATCTCAG 2628  
QY 2581 AGCAAGAGTCCCAAGCATGGGTGAACCTTGTGAGAGAGCTGACAGCTGCTGCTGCTGCT 2640  
Db 2629 AGCAAGAGTCCCAAGCATGGGTGAACCTTGTGAGAGAGCTGACAGCTGCTGCTGCTGCT 2688  
QY 2641 GTTCCAGAGCGAGGCTCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
Db 2689 GTTCCAGAGCGGAGGCTCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2748  
QY 2701 AGCAAGCGCTGGAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
Db 2749 AGCAAGCGCTGGAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2808  
QY 2761 CCCAGGAAGCTGCTCAGAGAGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2819  
Db 2809 CCCAGGAAGCTGCTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2868  
QY 2820 CTCTCTCAATTCCTCTC-ACTCTCTCAGAAACAGAGCTTGGCAGGGGAGAGAGGAGG 2878  
Db 2869 CCCAGCAATTCCTCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2928  
QY 2879 GCGCGGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2938  
Db 2929 GCGCGGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2988  
QY 2939 CGGTGACTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2998  
Db 2989 CGGTGACTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3048  
QY 2999 TGAAGCGTACGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3058  
Db 3049 TGAAGCGTACGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3108  
QY 3059 CCACATCCATGAGCCCAATGATCGGTGATCAATAA 3093  
Db 3109 CCACATCCATGAGCCCAATGATCGGTGATCAATAA 3143

RESULT 5  
US-09-957-187-5  
; Sequence 5, Application US/09957137  
; Publication No. US2003005414A1  
; GENERAL INFORMATION:

1 APPLICANT: Shimkels, Richard A.  
2 APPLICANT: LaRoche, William  
3 TITLE OF INVENTION: NOVEL POLYPEPTIDES AND PROTEINS ENCODED THEREBY  
4 FILE REFERENCE: 15966-540 CIP  
5 CURRENT APPLICATION NUMBER: US/09/957,187  
6 CURRENT FILING DATE: 2000-09-19  
7 PRIOR APPLICATION NUMBER: 60/123,667  
8 PRIOR FILING DATE: 1999-03-09  
9 PRIOR APPLICATION NUMBER: 09/520,781  
10 PRIOR FILING DATE: 2000-03-03  
11 PRIOR APPLICATION NUMBER: 60/234,082  
12 PRIOR FILING DATE: 2000-09-20  
13 PRIOR APPLICATION NUMBER: 60/233,798  
14 PRIOR FILING DATE: 2000-09-19  
15 PRIOR APPLICATION NUMBER: 60/174,485  
16 PRIOR FILING DATE: 2000-01-04  
17 NUMBER OF SEQ ID NOS: 85  
18 SOFTWARE: PatentIn Ver. 2.1  
19 SEQ ID NO 5  
20 LENGTH: 3333  
21 TYPE: DNA  
22 ORGANISM: Homo sapiens  
23 FEATURE:  
24 NAME/KEY: CDS  
25 LOCATION: (214)..(2865)  
26 NAME/KEY: misc.feature  
27 LOCATION: (2882)  
28 OTHER INFORMATION: an n may be any one of a or t or g or c  
29 US-09-957-187-5

Query Match 87.2%; Score 2698; DH 11; Length 3333;  
Host Local Similarity 93.9%; Pred. No. 0;  
Matches 2507; Conservative 0; Mismatches 21; Indels 167; Gaps 3;

QY 1 ATGAGTCAGAGGCTGCTGCTATATTTCACACTGCTACACTTGGCTGGGCTGGTTC 60  
DB 214 ATGAGTCAGAGGCTGCTGCTATATTTCACACTGCTACACTTGGCTGGGCTGGTTC 273  
QY 61 CCAGAAATCTGAGCAATCAGTATTGCTATGCAAGCACTATACAAACAGTATCCGGT 120  
DB 274 CCAGAAATCTGAGCAATCAGTATTGCTATGCAAGCACTATACAAACAGTATCCGGT 333  
QY 121 TTGTGGCCACAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATG 180  
DB 334 TTGTGGCCACAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATG 393  
QY 181 ATTATGATCATGAACGGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 394 ATTATGATCATGAACGGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453  
QY 241 ATAGACATCATACACAGGAGAGAAATTTATTTAGCAAAACATGACATGGAATCTAGA 300  
DB 454 ATAGACATCATACAGGAGAGAAATTTATTTAGCAAAACATGACATGGAATCTAGA 513  
QY 301 CAGGCCGATGATGACACATGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
DB 514 CAGGCCGATGATGACACATGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 573  
QY 361 ATAAAGTCTTCTAAAGAAACGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 574 ATAAAGTCTTCTAAAGAAACGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633  
QY 421 AACCTTCTCGAGAACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
DB 634 AACCTTCTCGAGAACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 693  
QY 481 GGAATGGCAGATGCCCATATGATGCCAAACAGTGGCACTGCTGCTGCTGCTGCTGCTGCT 540  
DB 694 GGAATGGCAGATGCCCATATGATGCCAAACAGTGGCACTGCTGCTGCTGCTGCTGCTGCT 753  
QY 541 AAATATATCTAGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 600  
DB 754 AAATATATCTAGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 813



DB 2809 CCCAGCAACCTGCTCAGAGAGAGCCACCTCAGCACTACTCTCATCAGAGCAACACTAA 2868  
QY 2820 CTCCTCAATTCCTTCIC-ACCTCTCCAGAACAGAGACTTTCAGAGGAGAGCAACCCGG 2878  
DB 2869 CCCGACAAATTCANCTCIGATTCARAGGAGCCAGAGCTTGGCAGGAGAGACACCCGG 2925  
QY 2879 CGCCGCCCCCGCAGAGGGTGACTCCATCCAGTGGACAGCTCCAGGCCATCTGGCCAGG 2939  
DB 2929 CGCCGCCCCCGCAGAGGGTGACTCCATCCAGTGGACAGCTCCAGGCCATCTGGCCAGG 2988  
QY 2939 CGGTGACTGTCTCGAGGAGTCCAGAGCTCAAGCGCTACAACTCACTGACAGAGTGGGGC 2998  
DB 2989 CGGTGACTGTCTCGAGGAGTCCAGAGCTCAAGCGCTCAAGCGCTACAACTCACTGACAGAGTGGGGC 3048  
QY 2999 TGAAGGTACGCCCTTCGCTAAAGCGGAGCTAGCCGCCCAAGCACTGCTTTCCTCCCTTT 3056  
DB 3049 TGAAGGTACGCCCTTCGCTAAAGCGGAGCTAGCCGCCCAAGCACTGCTTTCCTCCCTTT 3108  
QY 3059 CCAGATCCATGAAGCCCAATGATGCTGTACATAA 3094  
DB 3109 CCAGATCCATGAAGCCCAATGATGCTGTACATAA 3143

RESULT 6  
US-09-991-053-29  
Sequence 29, Application: US/0999-054  
Publication No. US2003003532A1  
GENERAL INFORMATION:  
APPLICANT: Schinkets, Richard A.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES ENCODING HUMAN SLIT-1  
TITLE OF INVENTION: MESEF-1 AND ROUNDABOUT-LIKE P-GLYCOPROTEIN  
FILE REFERENCE: 15966-540 CON S-10  
CURRENT APPLICATION NUMBER: US/09/391-053  
CURRENT FILING DATE: 2002-05-23  
PRIOR APPLICATION NUMBER: USN 63/123-657  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/525,781  
PRIOR FILING DATE: 2000-03-08  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 29  
LENGTH: 1890  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1890)  
US-09-991-053-29

Query Match: 60.8%, Score 1881.6, DR 11, Length 1890:  
Best Local Similarity 59.8%, Pred. No. 0:  
Matches 1884; Conservative 0; Mismatches 4; Indels 0; Gaps 0:

QY 49 GGGCTGGTTTCCCAAGAGATCTTGAGCAATAGTATTCGATGCAATATACAAAA 108  
DB 1 GGTATCCGGTTTCCCGAGAGATCTTGAGCAATAGTATTCGATGCAATATACAAAA 60  
QY 109 CAGTATCCGGTTTGTGGGCGCAGAGCCAGGAGGAGCAACACCAACAGAGACAGGCTG 168  
DB 61 CAGTATCCGGTTTGTGGGCGCAGAGCCAGGAGGAGCAACACCAACAGAGACAGGCTG 120  
QY 159 GACATCCAGATGATTATCATGATGATGATGATGATGATGATGATGATGATGATGAT 228  
DB 121 GACATCCAGATGATTATCATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 229 TATACCTGTTGATATAGACATATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 288  
DB 181 TATACCTGTTGATATAGACATATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
QY 289 TGGAAATCTGACAGGCGGATGTAGACATATGAGATGAGATGAGATGAGATGAGATGAG 348  
DB 241 TGGAAATCTAGACAGGCGGATGTAGACATATGAGATGAGATGAGATGAGATGAGATGAG 300

QY 349 TGCACAACTTATTAAGTTCTTCTAAAGAAACAGATCAATGATGATGATGATGATGATGAT 408  
DB 301 TGCACAACTTATTAAGTTCTTCTAAAGAAACAGATCAATGATGATGATGATGATGATGAT 360  
QY 409 ACTATGCTTCACACCTTCTCAGAAACTATAAGATGATGATGATGATGATGATGATGATGAT 468  
DB 361 ACTATGCTTCACACCTTCTCAGAAACTATAAGATGATGATGATGATGATGATGATGATGAT 420  
QY 469 GATGAATTCACGGAATGGCCAGATGCCATATGATGCCAAACATGCCAAACGTTGCACTG 528  
DB 421 GATGAATTCACGGAATGGCCAGATGCCATATGATGCCAAACATGCCAAACGTTGCACTG 480  
QY 529 TTTGAGATGAGAAACTATACTGAGCCAGTACTGACTGACTGACTGACTGACTGACTGACT 588  
DB 481 TTTGAGATGAGAAACTATACTGAGCCAGTACTGACTGACTGACTGACTGACTGACTGACT 540  
QY 589 ATTACCCGAGTCTTGAGAAAGCCCTACCTGCGGACCGTCAAGCAGCAGATTCAGAAATGG 648  
DB 541 ATTACCCGAGTCTTGAGAAAGCCCTACCTGCGGACCGTCAAGCAGCAGATTCAGAAATGG 600  
QY 649 TTGAAGAACCACTACTTTGTTCAAGCCGTGAGTACGAGAGTATATATCTACTTCTTCTTC 708  
DB 601 TTGAAGAACCACTACTTTGTTCAAGCCGTGAGTACGAGAGTATATATCTACTTCTTCTTC 660  
QY 709 AGGGAATAGCAGTGGAGTATACCAACAGGAAAGTAGTTTCCCAAGAGTGGGTGAG 768  
DB 661 AGGGAATAGCAGTGGAGTATACCAACAGGAAAGTAGTTTCCCAAGAGTGGGTGAG 720  
QY 769 GTTGTAGATGATGATGAGGAGTCTCAAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTG 828  
DB 721 GTTGTAGATGATGATGAGGAGTCTCAAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTG 780  
QY 829 CTGAGGCGGCTTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 888  
DB 781 CTGAGGCGGCTTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 840  
QY 889 CAGGAGTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948  
DB 841 CAGGAGTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 949 ATACCTTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008  
DB 901 ATACCTTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 1009 AGTGTGTTTACTGAGAGTCTCAAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1068  
DB 961 AGTGTGTTTACTGAGAGTCTCAAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1020  
QY 1069 COTGATGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1128  
DB 1021 COTGATGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1080  
QY 1129 AGATATGCAACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188  
DB 1081 AGATATGCAACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
QY 1189 CTCATGATGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1248  
DB 1241 CTCATGATGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1200  
QY 1249 AGATATGCAACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308  
DB 1201 AGATATGCAACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 1309 GTGGTTTCTGGGATCAGAGAGGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1368  
DB 1261 GTGGTTTCTGGGATCAGAGAGGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 1320  
QY 1369 AGTGTGTTTCTAAATGACAGCTTCTGAGAGGAGTACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1428  
DB 1321 AGTGTGTTTCTAAATGACAGCTTCTGAGAGGAGTACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1380

QY 1422 TCGAGCTATGATGGAGTGGAGAGAGAAAGCAATATGAGGATGAGCTGAGCTGAGAGAGAGAG 1428  
DB 1381 TCGAGCTATGATGGAGTGGAGAGAGAAAGCAATATGAGGATGAGCTGAGCTGAGAGAGAGAG 1440  
QY 1455 AGCTCTCTATGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1456  
DB 1441 AGCTCTCTATGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1500  
QY 1549 CGACATGGAGAGTGTAAAGAAAGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1609  
DB 1501 CGACATGGAGAGTGTAAAGAAAGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1560  
QY 1609 AAGGAGTGTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1668  
DB 1561 AAGGAGTGTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1520  
QY 1669 ATAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1728  
DB 1621 ATAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1680  
QY 1729 GGGCATTCAGTTCCTCTTCCGAGCAGACACACACACACACACACACACACACACACACAC 1788  
DB 1681 GGGCATTCAGTTCCTCTTCCGAGCAGACACACACACACACACACACACACACACACACAC 1740  
QY 1789 TATGAGTCTAGGGAGGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1848  
DB 1741 TATGAGTCTAGGGAGGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1800  
QY 1849 GAGCTTGTGGGGAGTCTTCCCATTAATCAGCAGACAGAGAGGAGTGTGAGCTGAGCTGAG 1908  
DB 1801 GAGCTTGTGGGGAGTCTTCCCATTAATCAGCAGACAGAGAGGAGTGTGAGCTGAGCTGAGCT 1860  
QY 1909 AGTACTTCAAGGCCAGCAGCAGCTGG 1936  
DB 1861 AGTACTTCAAGGCCAGCAGCAGCTGG 1886

## RESULT 7

US-09-957-187-29  
Sequence 29, Application US/09951187  
Publication No. US2003005414A1  
GENERAL INFORMATION:  
APPLICANT: Shimkets, Richard A.  
APPLICANT: LaRoche, William  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY  
FILE REFERENCE: 15466-545 CIP  
CURRENT APPLICATION NUMBER: US/09/957,187  
CURRENT FILING DATE: 2000-03-19  
PRIOR APPLICATION NUMBER: 60/123,667  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/520,781  
PRIOR FILING DATE: 2000-03-23  
PRIOR APPLICATION NUMBER: 60/234,082  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 60/233,796  
PRIOR FILING DATE: 2000-08-19  
PRIOR APPLICATION NUMBER: 60/174,485  
PRIOR FILING DATE: 2000-01-04  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentLis Ver. 2.1  
SEQ ID NO 29  
LENGTH: 1890  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1890)

## US-09-957-187-29

Query Match: 60.8%; Score 1881.6; DB 1; Length 1890;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1884; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 49 GGGGATGGTTCCTCCAGAGAGATTCAGGCAATCATAGTATTCGCATGCGCACTATACAAAA 108  
DB 1 GGAATCGGCTTCCTCCAGAGAGATTCAGGCAATCATAGTATTCGCATGCGCACTATACAAAA 60  
QY 109 CAGTATCGGCTTCCTCCAGAGAGATTCAGGCAATCATAGTATTCGCATGCGCACTATACAAAA 168  
DB 61 CAGTATCGGCTTCCTCCAGAGAGATTCAGGCAATCATAGTATTCGCATGCGCACTATACAAAA 120  
QY 169 GACATCCAGATGATATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228  
DB 121 GACAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 229 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 286  
DB 181 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
QY 289 TGGAAATCTAGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 348  
DB 241 TGGAAATCTAGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 349 TGGCAAACTTTATTAAGATTCCTTAAGAAAAAGATGATGATGATGATGATGATGATGATGAT 408  
DB 301 TGGCAAACTTTATTAAGATTCCTTAAGAAAAAGATGATGATGATGATGATGATGATGATGAT 360  
QY 409 ACTAATGCTTCAACGCTTCCTGAGAGAACTATAGATGATGATGATGATGATGATGATGATGAT 468  
DB 361 ACTAATGCTTCAACGCTTCCTGAGAGAACTATAGATGATGATGATGATGATGATGATGATGAT 420  
QY 469 GATGAATTCAGCGGAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528  
DB 421 GATGAATTCAGCGGAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 529 TTGCGAGATGGAAGAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588  
DB 481 TTGCGAGATGGAAGAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 589 AATTACCGGAGTCTTGAGAGAAAGCTTACCTGCGGAGCGTCAAGCAGGATTCAAAATGG 648  
DB 541 AATTACCGGAGTCTTGAGAGAAAGCTTACCTGCGGAGCGTCAAGCAGGATTCAAAATGG 600  
QY 649 TTGAAGAAGACCATGATTTCTCAAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 708  
DB 601 TTGAAGAAGACCATGATTTCTCAAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 709 AGGGAATAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768  
DB 661 AGGGAATAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
QY 769 GTTGTAGAATGATATGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828  
DB 721 GTTGTAGAATGATATGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 829 CTGAAGCGCGCTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888  
DB 781 CTGAAGCGCGCTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 889 CAGGAGTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948  
DB 841 CAGGAGTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 949 ACACCTTATAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008  
DB 901 ACACCTTATAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 1009 AGTGTGTTTACTGGGAGATTCAGGAGAGAGATCTGCTGATTCACCTGGACACCACTT 1068  
DB 961 AGTGTGTTTACTGGGAGATTCAGGAGAGAGATCTGCTGATTCACCTGGACACCACTT 1020  
QY 1069 CTTGATGAGAGAGTCTTAAAGTCCAGCGGAGGATGCTGCTGCTGGCTCATCTCTCTTAGAA 1126  
DB 1021 CTTGATGAGAGAGTCTTAAAGTCCAGCGGAGGATGCTGCTGCTGGCTCATCTCTCTTAGAA 1080  
QY 1129 AGATATGCAACCTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188







Sequence 60, Application US/09925299  
Patent No. US20020055627A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 60  
LENGTH: 1472  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (129)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc\_feature  
LOCATION: (130)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-60

Query Match 18.38: Score 565; DB 9: Length 1472;  
Best Local Similarity 99.5%; Prod. No. 3.4e-166;  
Matches 576; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 2515 GCCAGATGGCGTGGAGGACCGAGCGGCCGACACTGAGTATAGACCACTCAAGGAACAT 2574  
DB 13 GCCAGATGGCGTGGAGGACCGAGCGGCCGACACTGAGTATAGACCACTCAAGGAACAT 72  
QY 2575 CTGACGAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTGGCC 2634  
DB 73 CTGACGAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTGGCC 132  
QY 2635 CCAAAAGTTCCACAGCGGAGCGGCTGCTGGGTGAGGAGGAGGAGGAGGAGGAGGAGG 2694  
DB 133 CCAAAAGTTCCACAGCGGAGCGGCTGCTGGGTGAGGAGGAGGAGGAGGAGGAGGAGG 192  
QY 2695 GGCTTAAGCAAGCGCTGGAAATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2754  
DB 193 GGCTTAAGCAAGCGCTGGAAATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252  
QY 2755 AGTACCCCGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2814  
DB 253 AGTACCCCGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 311  
QY 2815 ACTAACTCTCCAAATTCCTCTCAGGCTCTCCAGAAACAGAGGCTTGGAGGGGAGAGAAC 2874  
DB 312 ACTAACTCTCCAAATTCCTCTCAGGCTCTCCAGAAACAGAGGCTTGGAGGGGAGAGAAC 371  
QY 2875 CCGCGCGCGCGCGCGAGAGGGTGGATCCATCCAGTGGACAGCTCCAGGAGGAGGAGGAGG 2934  
DB 372 CCGCGCGCGCGCGCGAGAGGGTGGATCCATCCAGTGGACAGCTCCAGGAGGAGGAGGAGG 431  
QY 2935 CAGGCGGTGACTGTCTCGAGGAGCGGAGGCTCAAGCGCTTACAACTCACTGACAAGGTGG 2994  
DB 432 CAGGCGGTGACTGTCTCGAGGAGCGGAGGCTCAAGCGCTTACAACTCACTGACAAGGTGG 491  
QY 2995 GGGCTGAAGCGTACGGCTTCGCTAAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 3054  
DB 492 GGGCTGAAGCGTACGGCTTCGCTAAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 551  
QY 3055 CTCTCCACATCCATGAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 3093  
DB 552 CTCTCCACATCCATGAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 590

RESULT 11

US-09-925-299-50  
Sequence 60, Application US/09925299  
Patent No. US20030340617A9  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 60  
LENGTH: 1472  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (129)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc\_feature  
LOCATION: (130)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-60

Query Match 18.38: Score 565; DB 11: Length 1472;  
Best Local Similarity 99.5%; Prod. No. 3.4e-166;  
Matches 576; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 2515 GCCAGATGGCGTGGAGGACCGAGCGGCCGACACTGAGTATAGACCACTCAAGGAACAT 2574  
DB 13 GCCAGATGGCGTGGAGGACCGAGCGGCCGACACTGAGTATAGACCACTCAAGGAACAT 72  
QY 2575 CTGACGAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTGGCC 2634  
DB 73 CTGACGAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTGGCC 132  
QY 2635 CCAAAAGTTCCACAGCGGAGCGGCTGCTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2694  
DB 133 CCAAAAGTTCCACAGCGGAGCGGCTGCTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 192  
QY 2695 GGCTTAAGCAAGCGCTGGAAATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2754  
DB 193 GGCTTAAGCAAGCGCTGGAAATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252  
QY 2755 AGTACCCCGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2814  
DB 253 AGTACCCCGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 311  
QY 2815 ACTAACTCTCCAAATTCCTCTCAGGCTCTCCAGAAACAGAGGCTTGGAGGGGAGAGAAC 2874  
DB 312 ACTAACTCTCCAAATTCCTCTCAGGCTCTCCAGAAACAGAGGCTTGGAGGGGAGAGAAC 371  
QY 2875 CCGCGCGCGCGCGCGAGAGGGTGGATCCATCCAGTGGACAGCTCCAGGAGGAGGAGGAGG 2934  
DB 372 CCGCGCGCGCGCGCGAGAGGGTGGATCCATCCAGTGGACAGCTCCAGGAGGAGGAGGAGG 431  
QY 2935 CAGGCGGTGACTGTCTCGAGGAGCGGAGGCTCAAGCGCTTACAACTCACTGACAAGGTGG 2994  
DB 432 CAGGCGGTGACTGTCTCGAGGAGCGGAGGCTCAAGCGCTTACAACTCACTGACAAGGTGG 491  
QY 2995 GGGCTGAAGCGTACGGCTTCGCTAAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 3054  
DB 492 GGGCTGAAGCGTACGGCTTCGCTAAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 551  
QY 3055 CTCTCCACATCCATGAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 3093  
DB 552 CTCTCCACATCCATGAAGCGGAGGAGTACCGCGCCAAACATCTTGTCTCC 590





QY	73	GAGCGCAATCAATATTGGUATGGCAATATACAAACAGATATCGGCTGTGTTGGGCGAC	132
Db	112	CGCGGGTTACCGCTGGCCCAAGAGAACTATGTGAACAACTATGCTGGCTGTGCTGGAGC	171
QY	133	AAGCGACGAGCAACAGCA-----ACAGAGACAGAGCGCTGGACATGATGATTATG	186
Db	172	GGCGCGGACCGCTGACCGCGGCGAGAGAGTGTGAGAGACTGCAATCAGCGCGAGTCTGTG	231
QY	187	ATCATGAGGAGGACCCCTACATCTGATAGAGTGTATATTAATATGTTGATATAGAG	246
Db	232	CGGCTGAACAGGAGCTGTCTCATCTGGGACAGAGGAGAGCTGTACGGGAGAGTGTGAT	291
QY	247	ACATCAGACACGGAGAGAAATTATCTGACAAAGACTGCAATGAAATATAGAAGAGC	306
Db	292	CCCGCCAGCTCCAGGAGCTGGGCTACAGAGAGAGCGAGCTGGAGATGTAAAGGACAG	351
QY	307	GATGTAGACATGCAAGATGAAGGAGAAACATAGAGATGATGGCAACATCTATTAAA	356
Db	352	GACATAAAGGTGTGCGGATGAAGGCAACAGGAGCGGAGCTGCGAATCTGGTAAG	411
QY	367	GTCTCTTAAAGAAACGATGATGATTTGTTGTCTGTGGAACTAATGCGCTCAACGCT	426
Db	412	GTGCTGTCTTGGGAGGAGTCCAGCGCTCTTTGTGTGCGGTTCCAGCGCTCAACCG	471
QY	427	TGCTGCAGAAACTATAGATCGATACATATGGAGCAATTGGGATTAATTCACGGGAATG	486
Db	472	GTGTCGGCAACTACGATAGACAGCACTGGACGCGTGGAGACACATCAGCGGATG	531
QY	487	GCGAGATGCCATATGATGCGCAACATAGCAAGCTTGTCACATGTTTGAGATGGAAACTA	546
Db	532	CGCGCGTGGCGGTACGACCGCAAGCAGCGCAATGTGGCTCTCTCTGACGGGATGCT	591
QY	547	TACTCAGCCACAGTGACTGTCCCTTGGCAATGACGAGTGCAITTACCGGAGTCTTGG	606
Db	592	ITCACAGTACTGTACCGACTTCCTAGCCATTGATGCTGTCACTACGCGAGCTGG	651
QY	607	GAAGCCCTACCTTGGGAGCGGTACAGCAGGATTCMAAATGGTTGAAGACACCAATCTT	666
Db	652	GACAGGCGACCGCTGGCACTGTGAACATGACTCCAAAGTGGTTCAAAGAGCCCTAGCTT	711
QY	667	GTTCAGCGGTGGATTACGCAGATTATATCTACTCTGTGTCAAGGAAATAGCAGTGGAG	726
Db	712	GUCCATCGGTGGAGTGGCGACGCAATGTACTCTCTGTGTGAGAGATTCGGATGGAG	771
QY	727	TATAACCATATGAAAGATAGTATTTTCCAAAGATGGTTCAGGTTTGTAAAGATATATG	786
Db	772	TTTAACTACTGGAGAAAGTGGTGGTGTGGCGGTGGGAGAGTGTGCAAGAGCGGCTG	831
QY	787	GGACATCTCAAGAGTCTCGAGAAATAGTGGAGGTGTCTTCAAGGCGGCTTCAAC	846
Db	832	GGAGCTTCCCGCGTCTGTGGAGAAATAGTGGAGTGTCTTCAAGGCGGCTTCAAC	891
QY	847	TGCTCACTTCTGGAGACTTCATTTTATTTCAACATCTCCAGGCGAGTTACAGATG	906
Db	892	TGCTCTGTACCCGGAGAGTCCCATTTCTATTTCAAGTGTTCAGGCTGTCAAGAGCGT	961
QY	907	ATTGCTATCAGGGCGTGAATGTGCTGGGAGAGCTTTTATACACTTATATAGCATC	966
Db	952	GTCAACCTCGGGCGCGCGGTGTCTGTGGCGGTTTCTTCAGGCGGAGCAAGCATC	1011
QY	967	CGTGGCTGGAGCTGTGGCTATGACATGCTTGACATGACAGTGTCTTACTGGGAGA	1026
Db	1012	CCTGCTCGGTGTCTGAGCTTTGACGTGGACAGTATGAGACTGTGTTCAGGCTGCG	1071
QY	1027	TTCAAGGACAGAACTCTCTCTGATTTGCTGCTGGACACCACTTCTGATGAACGAGTCT	1086
Db	1072	TCCGAGACAGAAAGTCCGCGGAGTCTCATCTGTGACGCTGTGGGAGGATCAGTGGCT	1141
QY	1087	AAGGCCAGGCGAGCTTCTGTGCTGCTCATCTCTGTACAAAGATATGCCAGCTGAT	1146
Db	1132	CGACCCCGGCGCGGTGCT-----CGGAGCGCGGTGATGCAATACAAAGCTCAGC	1191
QY	1147	GAGTTCCCTGATGATACCTTGAACTTCATCAGACGCGACCGCTCATGGATGAGCGAGT	1206

```

1186 GCCTTCGCGGATGACA-CCCTCAACTTTGTCAAGACCCACCCTCTGATGACGACGGCGGTG 1245
1207 CCCTCCATCTTCAACAGGCGCATGGTCTCGAGAACAAATGGTCAGATACCGCTTACCAA 1266
1246 CCCTCGCTGGGCCAATGGCCCTCGATCTCGATCTCGGACCCCTGATCAGGACCACTGCTACTCGA 1305
1267 ATGCGAGTGGACACATGTCGTGGGCCATATCAGAATCACACATGCTGTTTCTGGGATCA 1326
1306 CTGGCTGTGACAGTGGAGAGCGGCGCGCTGGGGCAACGACAGCCGTTGTCTTCCTGGGTCT 1365
1327 GAGAGGGAATCATCTTGAASCTTTT-----CGCCAGAA:AGGAATTAAGTGTTCCTA 1380
1346 GAGGCGGGGACGGTCTCTCAAACTTCTCTGTCGCGGCCCCAA:GCCAGCNCCTCAGGGACGCTCT 1425
1381 AATGACAGCCCTTTCCTGGGAGGATGAGTGTTTTCAACTCTGAATAATGGAG-----C 1434
1426 CGGCTCACTGTCTTCTCTGGAGGAGTTTGACACCTACCGGCGGACAGGTGTGGACGGCC 1485
1435 TATGATCGATCTCAAGACAAAGATCATGGGCATCGACCTGGACAGACAGCAGCAGCTCT 1494
1486 GCGGTCGCGAGACAGGCGACGGCGCTGAGCTGGAGCTGGACGCGAGCTTCGGGGGGC 1545
1495 CTGTATGTTGGCTTCTCTTACTGTCTG-CATAAAGTTTCCCTTGGCGGCTGTGAACGACAT 1554
1546 CTGCTGGCTGCTTCCCGCGCTCGGTGGTCCGAGTCTCTGGCTGCTGCCAGCAGTAC 1605
1555 GGGAGTGTAAANAACCTGTATTGCTCCAGAGACCCATATGTGGATGATGAATAAGGAA 1614
1606 TCGCGGTGTATGAGAAGTGTATCGGCGTCAAGACCCCTCTCGCGGTGGGCGCCCGGAC 1665
1615 GGTGTGCTCTCGACCCA:TTATCACCCCAACAGACAGACTGACTTTTTGACGAGCATAGAG 1674
1666 GCT--CCTGCATCTCTCTCAGCGCGGGCACAGAGCGCCTTTTGACGAGCAGTGTC 1722
1675 CTGGCAATACAGATGCTCTGGCGGACTG 1703
1723 CGGGCCAGCAGCTCAGCTTAGCGGACTG 1751

RESULT 14
US-10-035-977-34
: Sequence 34, Application US/10035977
: Publication NO. US20030134327A1
: GENERAL INFORMATION:
: APPLICANT: Desnoyers, Inc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Goddard, Aserey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Watanabe, Collin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS EXCEEDING THE SAME
: FILE REFERENCE: P3030R1C10
: CURRENT APPLICATION NUMBER: US/10/035,977
: CURRENT FILING DATE: 2001-12-26
: PRIOR APPLICATION NUMBER: 60/085579
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/112514
: PRIOR FILING DATE: 1998-12-15
: PRIOR APPLICATION NUMBER: 60/113300
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/113430
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/113605
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/113621
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/114140

```



Dd	652	GACAGGCCACCCCTGGGACACCTGTGAGACATGACCTCCAAATGGTTCAGAAAGGOUTTACTTT	711
Qy	667	GTTCAAGCGGIGATTTACGGAGATATATCTACTTCTTTCAGGGAAATAGCAATGGAG	726
Dd	712	GTCCATGGGTCGAGTGGGACGCGATGCTCTCTCTTCGGGAGACTGGATGGAG	771
Qy	727	TATACACCATGGAAGAGTACGTTTCCCAAGAGTGGCTCAGSTTTTCTGATGATGATG	786
Dd	772	TTTAATAGCTGGAGAAAGTGTGTGTGTGCTGTGGCTGTGGCTGATGTGCAAGACAGCTG	831
Qy	787	GGAGATATCTCAAGAGCTGTGGGAAGACATGTGAGCTGTTCTGAGAGGCTCTGAA	846
Dd	832	CGAGCTCTCCCGCGCTGCTGGAGAGCATGTGACGCTCTCTCTCAAGGCGGCGCTCAAC	891
Qy	847	TGCTCAGTCTCTGGAGACTCTCATTTTATTTCAAGACTTCTCAGGCAATTCAGATG	906
Dd	892	TGCTGTGACCGGAGACTCCATTTCTATCTTCAAGCTGTCTGAGCTCTCATGAGCTG	951
Qy	907	ATCTGATCAACGGGGGTGATTTTCTCTGTAAGCTTTCTACACTTATATACAGCATG	966
Dd	952	GTACAGCTCGGGGCGCGCGCTGTGCTGGCGCTTTTTCAGCGCCGACACAGCATG	1011
Qy	967	CTCGGCTGTGCAGCTGTGTGCTATGACATGCTGTGACATTTGCGCAGCTGTTTATG	1026
Dd	1012	CCTGGCTCGGCTGTCTCGCCCTTGACCTGACACAGGTGCGACCTGTGTTCGAAGCGCC	1071
Qy	1027	TTCAAGAGACAGAACTCTCTGATTCGACTGTGACGACCGAGTCTCTGAGAACAGTTCT	1086
Dd	1072	TTCCGAGGAGGAGTCTCCCGAGTCCATCTGACGCGCGGTGCGGAGGATCAGTGCCT	1131
Qy	1087	AAOCCGAGGCGAGTTGCTGTGTGCTCTATCTCTCTTGAAGAATATGCAACTCCAA	1146
Dd	1132	CGACCCCGGCGCGGTGCT-----GGCGAGCGCGCGGATTCAGTACAAATGCTCCAGC	1185
Qy	1147	GAGTTCCTCTGATGATCCGTGAACTTCATCAAGCAGCACCGCTCTGATGAGGCGATG	1206
Dd	1186	GCTTGGCGGATGACATCTTCAACTTTGTCAAGACCACCGCTCTGATGAGGAGGCGGTG	1245
Qy	1207	CGCTCCATCTTCAACAGGCGATGGTTCTTGAGAACAAATGGTCAGATACCGCTTACCAA	1266
Dd	1246	CCCTCGTGGCGATCGCGCTGATCTCTCGGAGCTGTATGAGGCGACGAGCTACTCGA	1305
Qy	1267	ATTGCACTGACACAGCTGCTGGCGATATCAGAAATCACAGTCTGGTTCCTGSGATCA	1326
Dd	1306	GTGGCTGTGACGTGGGAGCGGCGCTCGGAGCAAGCAGAGCTGTGCTCTCTGGGTCT	1365
Qy	1327	GAGAAGGGAATCATCTTGAAGTTT-----GGCGAATAAGAAATAGTGTGTTCTCA	1386
Dd	1366	GAGCGGGGACGGTCCCTCAAGTCTCTGTGTGGCGCAATGAGGACCTCAGGACCTCT	1425
Qy	1381	ATTATAGCTCTTCCGGAGGATGATGATTTTACACTCTTAAAAATGCAG-----C	1434
Dd	1426	GGGCTGAGTCTCTCTGGAGGAGTTTGAAGCTATAGCGCGACAGTGTGAGAGCGCC	1485
Qy	1435	TATGATCGAGTCAAGACAAAGAGCATGGGATGAGATGAGTGGACAGCAAGTNGCT	1494
Dd	1486	GGCGGTGGCGAGACGGCGAGCGCTCTGAGCTTGGAGTGGAGCGACTTGGGGGCG	1545
Qy	1495	CTGATGTTGGTTCTTACTGTGATTAAGTGTCCCTTGGCGGCTGTGAGCACT	1554
Dd	1546	CTGTGCTCTCCCTTCCCGGCTGGTGGTGGAGTGGCTTGGCTCTGCCACAGTAC	1605
Qy	1555	GGGAGTCTAANAAGCTGTATGCTCTGAGACAGCTATCTGTGATGCAATAGAGAA	1614
Dd	1606	TGGGGGTGTATGAAGAACTGTATGGGAGTTCAGACAGCGCTTACGCGGCTGGGCCG	1665
Qy	1615	GGTGTGCTCTGCAGCCATTATACCCCAAGCAGAGATGATTTGATTCAGACATAG	1674
Dd	1666	GGCT---CGTGCATCTTCTCTAGCCCGAGCACCCAGAACCGGCTTGGACAGGATGTCTC	1722
Qy	1675	CGTGGCAATACAGATGTTCTGGGGACTG	1703
Dd	1723	GGGCGCAGCACTCAGGCTTAGGGAGCTG	1753

[illegible]



GenCore version 5.1.1.6  
Copyright: (c) 1993 - 2003 Computer Mod.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:35:21 ; Search time 207.509 Seconds  
(without alignments)  
6578.995 Million cell updates/sec

Title: US-C9-856-631-1

Perfect score:

Sequence: 1 atagggtcagaagccttgc.....ccaatqatccqtatcataa 3593

Scoring table: IDENT- TV, NIC

Gapex 10.0 ; Gapext 1.0

Searched: 559978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 139956

Minimum DB seq length: 0

Maximum DB seq	Length: 2000000000
----------------	--------------------

.....

Post-processing: Minimum Match: 0%

Maximum Match 1008

Database : Issued\_Patents\_NA:\*

```
1: /cqn2_6/ptodata/i/ina/5A_COMB.seq;
```

2: /cqn2\_6/ptodata/1/ina/5B\_COMB.seq.\*

```
3: /cgn2_c/pidata/i/lna/6A_COMB.seq:
4: /cgn2_c/pidata/i/lna/6B_COMB.seq:
5: /cgn2_c/pidata/i/lna/6C_COMB.seq:
```

```
4: /cqn2_6/ptodata/1/lna/6B_COMB.seq:*
5: /cqn2_6/ptodata/1/lna/3CMTIS_COMB.seq:
```

```
3: /qnlz_6/products/ind/pcs_h.se
6: /sup2_6/products/ind/backfiles.se
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query		DB	ID	Description	
	Score	Match length				
1	555.4	18.0	3524	4	US-09-077-940A-3	Sequence 3, Appl
2	546.8	17.7	3692	4	US-09-077-940A-1	Sequence 5, Appl
3	443.6	14.3	2790	4	US-09-254-594-5	Sequence 4, Appl
4	443.6	14.3	3432	4	US-09-254-594-4	Sequence 2, Appl
5	369.8	12.0	2787	4	US-09-254-594-2	Sequence 1, Appl
6	369.8	12.0	3195	4	US-09-254-594-1	Sequence 57, Appl
7	165	5.4	2854	1	US-08-121-713D-57	Sequence 57, Appl
8	166	5.4	2854	1	US-08-835-268-57	Sequence 57, Appl
9	166	5.4	2854	2	US-09-060-692-57	Sequence 57, Appl
10	166	5.4	2854	3	US-09-833-391-57	Sequence 57, Appl
11	166	5.4	2854	4	US-09-060-610-57	Sequence 57, Appl
12	166	5.4	2854	5	PCT-US94-10151A-57	Sequence 57, Appl
13	146.4	4.7	3560	1	US-08-121-713D-59	Sequence 59, Appl
14	146.4	4.7	3560	1	US-08-835-268-59	Sequence 59, Appl
15	146.4	4.7	3560	2	US-09-060-692-59	Sequence 59, Appl
16	146.4	4.7	3560	3	US-08-833-391-59	Sequence 59, Appl
17	146.4	4.7	3560	4	US-09-060-610-59	Sequence 59, Appl
18	146.4	4.7	3560	5	PCT-US94-10151A-59	Sequence 59, Appl
19	142.2	4.6	2504	1	US-08-121-713D-63	Sequence 63, Appl
20	142.2	4.6	2504	1	US-08-835-268-63	Sequence 63, Appl
21	142.2	4.6	2504	2	US-09-060-692-63	Sequence 63, Appl
22	142.2	4.6	2504	3	US-08-833-391-63	Sequence 63, Appl
23	142.2	4.6	2504	4	US-09-060-610-63	Sequence 63, Appl
24	142.2	4.6	2504	5	PCT-US94-10151A-63	Sequence 63, Appl
25	126.8	4.1	2670	1	US-C8-121-713D-61	Sequence 61, Appl
26	126.8	4.1	2670	1	US-C8-835-268-61	Sequence 61, Appl
27	126.8	4.1	2670	2	US-09-060-692-61	Sequence 61, Appl

```
Db 312 CCCCCAGCTCCACGGAGCTCGGTACCAGGAGAGCTGACCTGGAGATCAACCCAGC 372
Qy 307 GATGTAGACATATGAGATGTAAGGAAACATATGAGTGAAGTCCACACCTTATATAA 366
Db 372 GACATAACGCTGTGTCGATGTAAGGGAACACAGAGGAGGAGTGTCCAAACTTCGTAAG 432
Qy 367 GTCTCTTAAGAAAAACGATGATGATGTTTGTCTGTGAAGTAATGCTTCAGCCCT 426
Db 432 GTCTGTCTCTTCGGGAGAGATCCACGCTCTTTGTGTGGGTTCGCAAGCCTTCAGCCG 492
Qy 427 TCCTGCAAAACTATAAGATCGATACATGGAACCATTCGAGGATGTAATTCAGCGAATG 486
Db 492 GTGTCCGCAACTACATACAGACACCTCGAGCCCTCGAGCAACATCAAGCGTATG 552
Qy 487 GCGAGATGCCCATATATGCCAATATGCGAGCTGCACTGTTGAGATGGAACCA 546
Db 552 GCGCGCTGCCGTACGACCTACGACAGCCCATGTTCGCTCTTCGTGAGGATGCTC 612
Qy 547 TACTCAGCCAGTACTGATTCCTTTGCCATTGACCCAGTCAATTTACCGGAGTCTTGA 606
Db 612 TTCAAGTACTGTATAGATCTTCATAGCATTCATGCTGATCTACCCAGGCTCGGG 672
Qy 607 GAAGGCTACCTGCGGAGCTCAAGTACGATTCAGAAATGTTGAAGAACATATCTT 666
Db 672 GACAGGCCACCTGCGCACCTGGAACATGACTCCAAAGTGGTTCAAGAGCTTACTT 732
Qy 667 GTTCAAGCCTGGATACGAGATATATATCTCTCTGAGTGAAGTGAAGTGGAG 726
Db 732 CTCTATGSGTGGATGAGGAGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 792
Qy 727 TATACAGCATGGAAGATGATTTCTCCAGAGATGCTCTCTCTCTCTCTCTCTCTCT 784
Db 792 TTTACTACTGAGAGATGATGCTGCTGAGAGATGCTCTCTCTCTCTCTCTCTCTCT 852
Qy 787 GAGCATCTCAAGAGTCTTCAAGAAACAGTGGAGTCTCTCTCTCTCTCTCTCTCTCT 846
Db 852 CGAGGCTCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912
Qy 847 TGCTCAGTCTCTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
Db 912 TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
Qy 907 ATTGATATACAGGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Db 972 GTCACCTCTGGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032
Qy 967 CTTGGGTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Db 1032 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
Qy 1027 TTCAGAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
Db 1092 TCCGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
Qy 1087 AAGCCAGGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
Db 1152 CGACCCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
Qy 1147 GAGTTCCTGATGATACCTTAACTTATCAAGAGGACCGGCTCATGGATGAGGAGT 1206
Db 1206 GCTTGGCGGATGATCTCTCACTTGTCAAGACCCACCTCTGATGAGGAGGCGGTG 1266
Qy 1207 GCTTCATCTTCAAGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
Db 1266 CCGCTGCTGCGGCGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
Qy 1267 ATGCGAGTGGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
Db 1326 GTGGCTGTGGAGTGGGAGCTGGGCGCTTGGGCGACATGAGACCGTGTCTCTCTCT 1386
Qy 1327 GAGAGGGAATCATCTGAAATTTT-----GGCGAAGTAAAGATGAGTCTCTCTA 1386
Db 1386 GAGGCGGGGAGGCTCTCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1446
```

## RESULT 2

```
US-09-077-940A-1
; Sequence 1, Application 95/09077940A
; Patent No. 6595442
; GENERAL INFORMATION:
; APPLICANT: KIMURA, TORU et al.
; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
; FILE REFERENCE: 0020-4426P
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent 1.0 version 3.1
; SEQ ID NO 1
; LENGTH: 3692
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(18)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (19)..(2682)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (2683)..(3653)
; OTHER INFORMATION:
; NAME/KEY: polyA_site
; LOCATION: (3654)..(3692)
; OTHER INFORMATION:
US-09-077-940A-1
```

Query Match 17.7%; Score 546.8; DB 4; Length 3692;  
Best Local Similarity 60.3%; Pred. No. 2.1e-136;  
Matches 1006; Conservative 0; Mismatches 632; Indels 30; Gaps 5;

```
Qy 57 TTTCGAGAGATCTTGAGCCATCAGTATTTCGCATGCGCACTATACAAACAGTATCC 116
Db 99 TTTCGAGAGATCAGCCACTCCACTCAGTGTGGCTCCAGGGAGTCTAGGCCACTACCC 158
Qy 117 GG-TGTGTGGGCGACCAAGCCAGGAGGAGCAACACACAGAGAGC-----ACAGGCTGA 170
Db 159 CGTGTTCGTGGGCGAGCGGCGCTGCTGCTGACCCCTGCAGAGGGTGTCTGAGGACCTCA 218
Qy 171 CATCCAGATGATTTATGATCATGACGAGACCCCTCTACATCTGTCTAGGAGACCATTTTA 230
Db 219 CATCCAGAGAGTCTACGTGTTTAAACAGGACACTGTTTCATCGGGGACAGAGACAACCTGA 278
```



QY 344 ATGAGTCCACACACTTAACTTCTTCTAAAGAAAGATGATGATCTTCTCT 403  
DQ 356 ATGAGTCTACAACTATATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 415  
QY 404 GTGGAACCTATGCTTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 463  
DQ 416 GTGGAACCTATGCTTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 475  
QY 464 TCGGGATGAATTCAGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 523  
DQ 476 AGGTGAGGAACTGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAG 535  
QY 524 CACTGTTGAGATGGAATACTATAGTCAAGCAAGTGAAGTCAAGTGAAG 583  
DQ 536 CCATCTTTGAGAGGAGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAG 595  
QY 584 CAGTCATTACCGGAGTCTTGGAGAAAGCTTACCGGAGTCTTACCGGAGT 643  
DQ 596 CTGTAGTTTACAGAGGCTTGGGCGGAGGCGGAGTCTTACCGGAGTCTT 655  
QY 644 AATGGTTGAAGACCATACCTTCTTCAAGCGTGTATACCGGAGTCTTAC 703  
DQ 656 AGTGGTCCGAGGCGGAGTCTTGGGCGGAGGCGGAGTCTTGGGCGGAG 715  
QY 704 TCTTCAGGGAATACAGTGTATACAGTGTATACAGTGTATACAGTGTAT 763  
DQ 716 TCTTCAGGGAATACAGTGTATACAGTGTATACAGTGTATACAGTGTAT 775  
QY 764 CTCAGGTTTGAAGATGATGCGGAGTGTATACAGTGTATACAGTGTATAC 823  
DQ 776 CCGGAGTATGAAGTGTATACAGTGTATACAGTGTATACAGTGTATACAG 835  
QY 824 CGTTCCTCAAGGCGGAGTGTATACAGTGTATACAGTGTATACAGTGTAT 883  
DQ 836 CCTTCCTCAAGGCGGAGTGTATACAGTGTATACAGTGTATACAGTGTAT 895  
QY 884 TTTCCAGGCGGAGTGTATACAGTGTATACAGTGTATACAGTGTATACAG 943  
DQ 896 TTTTCAGGCGGAGTGTATACAGTGTATACAGTGTATACAGTGTATACAG 955  
QY 944 TTTTCAGGCGGAGTGTATACAGTGTATACAGTGTATACAGTGTATACAG 1003  
DQ 956 TCACCAACGAGCAATAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1015  
QY 1004 TTGGCAGTGTATACAGGAGTGTATACAGGAGTGTATACAGGAGTGTAT 1063  
DQ 1016 TTGAGGCGGAGTGTATACAGGAGTGTATACAGGAGTGTATACAGGAGT 1075  
QY 1064 CAGTTCCTGATGAGGAGTGTATACAGGAGTGTATACAGGAGTGTATAC 1123  
DQ 1076 CTGTGCTGAGGAGGAGTGTATACAGGAGTGTATACAGGAGTGTATACAG 1135  
QY 1124 TGAAGAGATATGAGGAGTGTATACAGGAGTGTATACAGGAGTGTATACAG 1183  
DQ 1136 CTGCGTGTCT 1195  
QY 1184 ACCCGCTCATGATGAGGAGTGTATACAGGAGTGTATACAGGAGTGTATAC 1243  
DQ 1196 ACCCGCTCATGATGAGGAGTGTATACAGGAGTGTATACAGGAGTGTATAC 1252  
QY 1244 TGGTCAGATACCGGCTTACCAAAATTCAGTGTACAGGAGTGTATACAGG 1303  
DQ 1253 CTAGCAGGCGGCTTACGAGGAGTGTATACAGGAGTGTATACAGGAGTGTAT 1312  
QY 1304 ACACGTGTGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1363  
DQ 1313 TCACAGTCAITGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1372  
QY 1364 GAAATAGTGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1423  
DQ 1373 GCGGATCCGGTCT 1429  
QY 1424 AAAAATGAGGCTATGAGGAGTGTATACAGGAGTGTATACAGGAGTGTATAC 1477

## RESULT 4

US-09-254-594-4  
Sequence 4, Application US/09254594  
Patent No. 6566094  
GENERAL INFORMATION:  
APPLICANT: KIMURA, Toru  
INVENTOR: KIMURA, Toru  
TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y  
FILE REFERENCE: 0020-4527  
CURRENT APPLICATION NUMBER: US/09/254,594  
CURRENT FILING DATE: 1999-05-11  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 3432  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (3)..(187)  
OTHER INFORMATION: Tissue Type: Child Brain  
NAME/KEY: 5'UTR  
LOCATION: (1)..(187)  
OTHER INFORMATION: Identification Method: E  
NAME/KEY: misc\_feature  
LOCATION: (188)..(2977)  
OTHER INFORMATION: CDS: Identification Method: E  
NAME/KEY: 3'UTR  
LOCATION: (2978)..(3432)  
OTHER INFORMATION: Identification Method: E  
NAME/KEY: polyA\_signal  
LOCATION: (3408)..(3432)  
OTHER INFORMATION: Identification Method: E  
US-09-254-594-4

Query Match: 14.8%; Score 443.6; DB 4; Length 3432;

Best local similarity 58.6%; Pred. No. 9, 2e-109;

Matches 849; Conservative 0; Mismatches 584; Indels 15; Gaps 4;

QY 164 GCGTGGACATCCAGATGATTAATGATCATGACGAGGAGGAGGAGGAGGAGG 223  
DQ 366 GCGTGGATTCAGAGATGATTAATGATCATGACGAGGAGGAGGAGGAGGAGG 425  
QY 224 ATATATATATCTGATATAGACATCATACACGAGGAGGAGGAGGAGGAGGAGG 283  
DQ 426 AGTTCCT 485  
QY 284 TGACATGGAATCTAGACAGGCGGATGTAGACATGATGAGGAGGAGGAGGAGGAGG 343  
DQ 486 ATCTA---ACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542  
QY 344 ATGAGTCCACAACTTTATTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 403  
DQ 543 ATGAGTCTACAACTTTATTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 602  
QY 404 GTGGAAGTAAAGCTCTACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 463



D	593	CTGTGGTTTACAGAAAGCCTTGGACCTCAACCCGCCATCTGGTCTCTGCAAAAGTATGACATGCA	652
Q	644	AATGTTTGAAGAACCATACTTTGTTTCAAGCCGTGGATACAGGAGATTTATATCTACTTCT	703
D	653	AGTGGCTTCGAGAGCCACACTTGTCTATGCTTTGGAGCATGAGACCAATGTCACCTCT	712
Q	704	TCTTCAGGGAATAGCACTGGAGTATA--ACNCCATGGGAAAGTAGTCTTTCCCAAGAG	750
D	713	TTCTTCCGGAGAACTCTCTGTGAGAGACCCCGGGCTGGGAGGTGGAGTCTTCCCGCG	772
Q	761	TGGCTCAGGTTTGAAGAATGCAATGCGAGGATCTCAAAATATCTCTGGAGAAACAGTGG	820
D	773	TGGCCCGGCTGTGTAACAGTGCATGGTGGCTCAACAAGGCGCTTGCATCGCCACTGG	832
Q	821	CGTCGTTCCGAAAGCGGCTTGACCTGCTCACTTCTGGAGACTCTCATTTTATTTCTCA	880
D	833	CATCTCTTCCTTAACTGAGCTCAACTGCTGGGTGCTTGGAGACTTCACTTCTTACTTTG	882
Q	881	ACATCTCTCCAGGCACTTACAAATGTGATTCGTATCAACGGCGGTGATGTTCTCTGGCAA	940
D	893	ATGCTTACACTGCTTAACCTGGCTGTGAACCTGCATGGGGGCTCTGGCTCTTTGGG	952
Q	941	CGTTTCTACACCTTATAACAGCATCCCTGGCTCTGACGCTCTGACCTATGACATGTTG	1000
D	953	TCTTACACTACAGACCAATAGCACTCTGGGTCTGCACTTGGGCTCTTACCTAGATG	1012
Q	1002	ACATCTCCAGTCTTTTCTAGGAGATCTCAAGAACATAAGCTCTCTGATTCACCATGGA	1060
D	1013	ACATCTGACGCTGGCTTCTAGGCGCAAGTTCAAGGAGCAGAGAGTCTGATGGGCGCTGGA	1072
Q	1062	CAGCAGTCTCTGTGAGAGCAATTCGTAATCCAGAGCAGCTTGCTGCTGCTCACTCT	1120
D	1073	CTCTGTGCTGTGAGAGCAAACTCCCTCAGCAGGCGCAGGCTCTGTGCGAGTCTGGGTG	1132
Q	1121	CGTTAGAAGATATGCAACCTCCCAATGAGTTCCTGATCATACCTGAACTTCAATGAGA	1180
D	1133	CAGCTGCCCTATCTCTCTCTCTCAAGAGCTGGCTGACGATCTCTGCTCTTCATCAAGG	1192
Q	1181	CGCACCCGCTCATGATGAGGACAGTGGCTGCCCTCCATCTTCAAGAGGCCAAGGTCTGTGASNA	1240
D	1193	CACACCCACTGCTGGATCCCGCTGTGCAACC---TGCCACCCATCACTCTCTCTCTCTC	1249
Q	1241	CAATGGTCAGATACGGCTTACCAAAATTCAGTGGAGACAGCTCTGGGCGCATATGAGA	1300
D	1250	TGACATGAGCGGCTCTACTGACCCAGGAGCTGTGCACTGTAAGCTGGCCGCCACACAA	1309
Q	1302	ATCCACAGCTGTTTCTGTGATGAGAGAGGAGATGATCTGAGCTTTTGGCGCAAA	1360
D	1310	ATACTACAGCTCTGTTCTCTTGGCTCCATATGAGGAGAGTGTGAGGTTGCTGACTGAG	1369
Q	1361	TAGGAAAATGTTGTTTCTATAATGACAGCTTTTCTGAGAGCAATGAGTCTTACAACT	1420
D	1370	GGGAGACA---GTCTCTGGAGCCGAGCTATCATATTGGAAGAGATTGATGCTATAGCC	1426
Q	1421	CTGAAAATGCACTATGATGG-----AGTCGAATACAAAGAGTATCGGCTATGGAGC	1474
D	1427	ATGCCCGGTGCATGGGAGCGGTCAACCGAGCTCTCTGAGAGATCATGAGCTTGGAGC	1486
Q	1475	TGGAGAGAGCAAGCAGCTCTCTGATCTGTGCGTCTGTCTATCTGTGTGATGAAGTTCCGC	1534
D	1487	TGCACACTGAGGGTCAAGAGCTTTTGTGGCCCTTCTGTGATGCGATCTTACCTCTCTC	1546
Q	1535	TGGCCCGGTGTGACGACATGGGAGCTGTAAATAACCTCTTATCGCTCCAGACAGCCAT	1594
D	1547	TACGCGGTGTGCCCGGATGGAGCATGTTCAGAGGAGCTGCGCTGCGCTTCTCTTGGACCT	1606
Q	1595	ATTGTGATGG : 605	
D	1607	ACTGTGGAIG 1617	

```

1 Sequence i. Application US/09254594
2 Patent No. 6566094
3 GENERAL INFORMATION:
4 APPLICANT: KIMURA, Toru
5 APPLICANT: KIKUCHI, Kazuo
6 TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
7 FILE REFERENCE: 0020-4527P
8 CURRENT APPLICATION NUMBER: US/09/254,594
9 CURRENT FILING DATE: 1999-05-11
10 NUMBER OF SEQ ID NOS: 13
11 SOFTWARE: Patentin version 3.0
12 SEQ ID NO 1
13 LENGTH: 3195
14 TYPE: DNA
15 ORGANISM: Rattus norvegicus
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: ()..( )
19 OTHER INFORMATION: Tissue Type: Brain
20 NAME/KEY: S'UTR
21 LOCATION: (1)..(50)
22 OTHER INFORMATION: Identification Method: E
23 NAME/KEY: misc_feature
24 LOCATION: (51)..(2837)
25 OTHER INFORMATION: CDS; Identification Method: F
26 NAME/KEY: 3'UTR
27 LOCATION: (2838)..(3195)
28 OTHER INFORMATION: Identification Method: E
29 US-09-254-594-1

```

Query Match	12.0%	Score 369.8;	DB 4;	Length 3155;
Best Local Similarity	55.9%	Pred. No. 5,46-89;		
Matches 811;	Conservative	0;	Mismatches 622;	Indels 18; Gaps 5
QY	164	GGCTGGACATTCAGATGATTATGATCATGAAGGAAACCTCTACATTCGTCTGCTAGGAACT	223	
DB	226	GGCTGGACCTTCAGAGATTCCTGACCTTGAACCGGACCTTGGCTGTGGCTGCCGGGATC	285	
QY	224	ATATTATATCTGTTTCATATACACATCATCACACGGGAAGAAATTTATTTAGCAAAAAC	283	
DB	286	AGTTTTTCCTTCGATCTTCAGGCCAAGAAGGGGAGGGCTGTGCTCCCAACAGT	345	
QY	284	TGCATGTGAATCTTAGCAGGCGGATCTAGACATCATCAANTGAAGGGGAAAAACAT	343	
DB	346	TTCTGACATCGGGAGCCGACGACATGGAGAAATGTGCTTCGGGGGAAAGCTGACGG	402	
QY	344	ATGAGTGGCACAACCTATTAAAGCTCTTCTTAAAGAAAAAGCAIGATGCAATGTTT	403	
DB	403	ACGAATGCIACANCTACATCTGCTTCTTCTTCCCTGGGACCTCCAGACATCCTTGCT	462	
QY	404	GTGGAACTAATGCTTCAAGCTTCTCTGACAAACTATATAGATGATATGTAATTTGAACCAI	463	
DB	463	GTGGAACTAATTTCTCTGACCTCTGTGTGTCACACTATGGGATAACATCTCTGCAACAGS	522	
QY	464	TGGGGATGAATTCACATCGAATGCGGACATGCCCATATGATGCCAACATGCCAAGTTG	523	
DB	523	AGGGTGAGGAGCTGAGTGGGCAAGCTGCACTGCCCTTTGATGCCATCCAGTCCACTGTGG	582	
QY	524	CACCTGTTCCAGTGGAAAGCTATACCTCAGCCAGTGTACTGACTTCCTTCCCATTTGACG	583	
DB	583	CCATCTCTGACAGGGTAGTTGTACTCAGCCAGCAGGACAGATTTCCAGGCCAGTGTG	642	
QY	584	CAGTCATTATCGGAGCTCTTGAGAAAGCCCTACGCTCGCGAGCGCTCAGCAGCATTCAA	643	
DB	643	CTGTGGTTTACAGAAGCTCTGACCTCAGCCCCACCTCCGTTCTGCAAAATGATGACTCCA	702	
QY	644	AATGCTTGAAGAACTATACATTGTTCAAGCCGTGGATTAGGAGATTATATCTACITCTT	703	
DB	703	AGTGGCTTCGAGGCCACACTTTGCTATGCTTTGGAGCATGGAGACCAITGCTACTTCT	762	
QY	704	CTCTCAGGGAATACCTAGTGGAGTATA---ACACCATGGGAAAGGTAGTTTTCCCAAGAG	760	
DB	763	TTCTTCCGCAAGTCTCTGTGGAGAGCGCCGCGCTTGGGGAGGTGCATTTTCCCGG	822	

```
QY 761 TGCGTCAAGTTCTAGAAAGATATGGAGGATCTCAAAAGACTCGGAGAAACAGTGA 420
Db 823 TGCGCGGCTGTGAATGATGAGATGGTGGCTCAGCAGGGCTGGATGCGCACTGA 882
QY 821 GTCGCTGCTGAGGAGCGCTTGAATGCTGAGTCTCTGGAGACTCTCATTTTATTICA 880
Db 863 CATCTTCTCTAGCTGAGGCTCAACTGCTGGTCTCTGGGACTCTACCTTCACTTIG 942
QY 881 ACATTCTTAGGACGTTAGAGATGTGATATATGAAGGAGCTGATGTCTCTGGCAA 940
Db 943 ATGCTTACAGTCTCTTAACCTGAGGCTTGAGAGCTGATGAGGCTGCTGGCTTTGAG 1002
QY 941 CGTTTCTACACTTATACAGATCTGCTGGCTGCACTGCTGCTGCTGCTGCTGCTG 1000
Db 1003 TCTTCACTACTAGACACATAGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1062
QY 1001 ACATTCTGAGTGTCTTACCTGAGATCTCAAGGACAGAGTCTCTGATCTCAGTGA 1060
Db 1063 ACATTGAGCTGCTTGAAGGCAAGTCTCAAGGACAGAGTCTCTGATGCTGCTGCTGA 1122
QY 1061 CACCACTTCTGATGAAGGATGCTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120
Db 1123 CTCTCTGCTGAGGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1182
QY 1121 CTTTGAAGATATGCAACTCTCAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180
Db 1183 CAGCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1242
QY 1181 GCGACCGCTGATGATGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1240
Db 1243 CACACCCACTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1299
QY 1241 CAATGCTGATACCTGCTTACCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1300
Db 1300 TGACTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1359
QY 1301 ATCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1360
Db 1360 ATACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1419
QY 1361 TAGGAATAGTGTCTTCTTAAATGACAGCTTTCTGCTGAGGAGATGATGCTTACA 1420
Db 1420 GGGGACA---GCTCTGGGACCGGAGCTTATCATATGGAAGAGATGATGCTTACAG 1476
QY 1421 CTGAAAAATGCACTATGATG---AGTGAAGACAAAGGATCATGGGCTGCTGCTG 1474
Db 1477 ATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1536
QY 1475 TGGACAGACAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1534
Db 1537 TGGACACTGAGGCTCAGAGCTTTTGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1596
QY 1535 TTGGCGGCTGTAAGCAATGAGAACTGTAAGAAACCTGATGCTTCTGCTGCTGCTG 1594
Db 1597 TGAGCCCTGCTGCGGCGCATGAGAGTGCAGAGGAGCTGCTGCTGCTGCTGCTGCTG 1656
QY 1595 ATGAGGAGGG 1605
Db 1657 ACTGTGGATGG 1667
```

## RESULT 7

```
US-08-121-713D-57
; Sequence 57, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Mathes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
```

```
NUMBR OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 248 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713D
FILING DATE: 13-SEP-1993
CLASSIFICATION: 5.14
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,527
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 2854 base pairs
TYPE: Nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 451..2640
US-08-121-713D-57
```

## Query Match

5.4% Score 166; DB 2; Length 2854;  
Best Local Similarity 53.0%; Pred. No. 2.2e-34;  
Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

```
QY 289 TGAATATCTAGACAGCGCATGTAGACATGTCAGAAATGAAGGAAACATGAAGATGAG 348
Db 709 TGGCCTCGTCAGTGGCCATCGGAGCTCTGCTACTCAAGGGGAGTCAGAGGACGAC 768
QY 349 TGGCAGACATTTAATAGTCTTCTTAAGAAAACGATGATGCTGTTGCTGTGA 408
Db 769 TGGCAGAACTACATCGGAGTCTGGGAAATTTGAGGATCAGCGGTACTCTCTCGGT 828
QY 409 ACTAATCCCTTCAAGCTTCTCTGAGAAATATAAGAT---GGATACATTTGGAACCATTC 465
Db 829 AGAAGCCCTATAAGCCACTATGTGGCCTACGCTCAAGGATGAGATTAITGTTGA 888
QY 466 GGGATGAAATCAGCGCAATGGCAGATGCCATATGATGCCAAACATGCCAAGCTGCA 525
Db 869 GAGAAAGATATGAGGAGAGGATTTGCGCCATTTGAGCCTGACCAACAGCAGCTGCA 948
QY 526 CTGTTTGAAGTGAAGATATATCTAGCAGACATGATGATCTGCTGCTGCTGCTGCTG 585
Db 949 ATATACATGAGGAGAAATGTATACAGACAGACATGGCAGACTTCTGGAAGTACCT 1008
QY 586 GTCAATTACCGAGCTCTGAGAAAGCCCTTACCTGCGGACCGGCAAGCAGATTCAAA 645
Db 1009 CTCATATACCGGCTGCTGAGTACATCACTGCGGAAAGCTATCTATCTCTCTCT 1053
QY 646 TGGTGAAGACCATACATCTGTCAGCGCTGATTCAGGAGATATATCTCTCTCTCT 705
Db 1054 CAATTAATCTCTTAACTTGTCAACAACTGGAGTACAATGATTTATATCTCTCT 1113
QY 706 TCCAGGAAATAGCAGTGGAGTATACACCATGCGAAAGCTAGTTTCCCAAGAGTGG 765
Db 1114 TTCGAGAGACTGCTGTGAGTACATCACTGCGGAAAGCTATCTATCTCAAGAGTGG 1173
```







APPLICANT: Matthews, David  
 APPLICANT: Bentley, David R.  
 APPLICANT: O'Connor, Timothy  
 TITLE OF INVENTION: The Semaphorin Gene Family  
 NUMBER OF SEQUENCES: 160  
 CORRESPONDENCE ADDRESS: 160  
 ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 Bush Street, Suite 3200  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/833,391  
 FILING DATE:  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/121,713  
 FILING DATE: 11-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman, Richard A.  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: B94-002-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415)343-4341  
 TELEFAX: (415) 343-4342  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 57:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2854 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 451..2640  
 CS-08-833-391-57

Query Match 5.4%; Score 166; DB 3; Length 2854;  
 Best Local Similarity 53.0%; Pred. No. 2,2e-34;  
 Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

QY 289 TGGAAATCTAGACAGGCCGATGTAGACACATGACAGATGACAGGAAACATAGGATAG 348  
 DB 709 TGGAAATCTAGACAGGCCGATGTAGACACATGACAGATGACAGGAAACATAGGATAG 768  
 QY 349 TGGAAATCTAGACAGGCCGATGTAGACACATGACAGATGACAGGAAACATAGGATAG 408  
 DB 769 TGGAAATCTAGACAGGCCGATGTAGACACATGACAGATGACAGGAAACATAGGATAG 828  
 QY 409 ACTAATGCTTGAACCTTCTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 468  
 DB 829 ACTAATGCTTGAACCTTCTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 888  
 QY 466 GGGGATGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 525  
 DB 889 GAGAGAGATGAGGAGAGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAG 948  
 QY 526 CTGTTTGCAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAG 585  
 DB 949 ATATACAGTGGGAGCAATTTATGAGAGAGATGAGGAGATGAGGAGATGAGGAGATG 1008  
 QY 586 GTGATTACGGAGTCTTGGAGAGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAG 645  
 DB 1009 GTGATTACGGAGTCTTGGAGAGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAG 1063  
 QY 646 TGGTGAAGAGAGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 705

DB 1054 CAATTAAATGCTTAACTTGTCAACACAAATGAGTAAATGATTTATATCTCTTC 1113  
 QY 706 TCCAGGGAATAGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 765  
 DB 1114 TCCAGGGAATAGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1173  
 QY 766 CAGGTGTGTAAAGATGATATGAGGAGATGATCAAGAGATGCTGGAGAGAGATGAGGAG 825  
 DB 1174 AGAGTCTGTAAACATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230  
 QY 826 TCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885  
 DB 1231 TTTTGAATACAGTCTGAACTGTTCCGTCGCTGAGATTTATCCATTTTACTTCAATGAA 1290  
 QY 886 CTCAGGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941  
 DB 1291 ATTGAGTCAACAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350  
 QY 942 TTTTCTACAGCTTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993  
 DB 1351 TACGGTGTCTTACAGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1410  
 QY 994 ATGCTTACAGTGTCCAGTGTCTTACAGTGTCTTACAGTGTCTTACAGTGTCTTACAGT 1053  
 DB 1411 ATGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470  
 QY 1054 ACTGACAGGAGTGTCTGATGAGAGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1113  
 DB 1471 AACTGTTGCCAGTGTCCAGGAGTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1526  
 QY 1114 TCATCTCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173  
 DB 1527 TTTTCTACAGCTTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586  
 QY 1174 ATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1233  
 DB 1567 GTAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626  
 QY 1234 CTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1277  
 DB 1627 ATTGAGTCAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1670

RESULT 11  
 US-09-060-610-57  
 Sequence 57, Application US/09050610  
 Patent No. 6344544  
 GENERAL INFORMATION:  
 APPLICANT: Goodman, Corey S.  
 APPLICANT: Kolodkin, Alex L.  
 APPLICANT: Matthews, David R.  
 APPLICANT: Bentley, David R.  
 APPLICANT: O'Connor, Timothy  
 TITLE OF INVENTION: The Semaphorin Gene Family  
 NUMBER OF SEQUENCES: 160  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 Bush Street, Suite 3200  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/060,610  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:





407 GCGTGGCCGCTGCAAAATGGCAAAAGTGGCGGATGATTC---GCAACGGT 463  
818 GGAGCGCTGCTGCAAAAGTGGCAAAAGTGGCGGATGATTC---GCAACGGT 877  
464 GGAGCGCTGCTGCAAAAGTGGCAAAAGTGGCGGATGATTC---GCAACGGT 523  
878 TCACATATTCGCAAGGAGTACAGATGATGAT-----TCATATCAAGGAGG 925  
524 TTAATGAATCCAAATGCGCAAGATGCGGAGGAGGAGTATGATGCAAGAGCGA 583  
926 ATGTGTGCTGGCAAGCTTCTACACCTTAACAGGAGGAGGAGTATGATGCAAG 965  
584 AACTGATCTACGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAGGAG 643  
986 CCAATGATGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGG 1045  
644 CTTTGGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 703  
1045 CTGATGATGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGG 1105  
704 TCAACTGCAAGTCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAG 763  
1106 GCGTGGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 1165  
764 GTCACAGGATTCG-----AGAGGCTTCCGGATCCGAGCAG 799  
1166 TGAATTCATCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 1225  
800 TGAATTCATCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 859  
1226 CATGCTGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 1277  
860 CGATTTTGGCGGAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 911

RESULT 14

US-08-835-268-59  
; Sequence 59: Application US/08/835268  
; Patent No. 5807826  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S  
; APPLICANT: Kolodkin, Alex L  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 200  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent'n Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/835,268  
; FILING DATE:  
; PRIORITY DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,713  
; FILING DATE: 13-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: F94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)343-4341

TELEFAX: (415) 343-4342  
TELEX:  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3560 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1953  
; US-08-835-268-59  
Query Match 4.7% Score 146.4; DB 1; Length 3560;  
Best Local Similarity 51.9%; Pred. No. 4.4e-29;  
Matches 494; Conservative 0; Mismatches 401; Indels 57; Gaps 5;  
QY 341 AGGATGAGGCGCAAGCTTATTAAAGTTCTTCAAGAAAGAGGATGATGATGTTG 400  
DB 2 AGGATGATGTCAGAAATACATCCGATCATGGTGGTGGCGGCGGCTGCGCTTTCG 61  
QY 401 TCTGTGAACATATGCTTCAACCTTCCCTGCAAGAACTATAAGAT---GATACATG 457  
DB 62 TTTGGGCAACAGCTGTTCCGGCCCATGCGACACGATATATCATATTAGTGACAGCA 121  
QY 458 AACCATTCGGGGAATGAAATTCAGCGGGAATGGCCAGATGCCATATGATGCAACAG 517  
DB 122 ACAGGCTGGAGGCGGCAAGAAAGGACAGGCTGGTGTGCGCTACGATCCAGCTCA 181  
QY 519 ACCTTGGACTGTTGCAGATGGAAGAACTATATACAGCCAGAGTACGATGCTTCC 577  
DB 132 CCACCTGCTGCTGGCGGACAGCACTGATTCGGGTACGGTGGCGGATTCAGTGGCA 241  
QY 578 TTGAGCGAGTCATTTACCGGAGCTTTGGAGAAAGCCCTACCCCTGGGAGCGTCAAG 637  
DB 242 GCGATCCCATATCTACCGGAGC-----CCCTGCAGACCGAGCAGTACG 286  
QY 638 ATTCAAATGTTGAAACAGCCATCTTGTCAAGCGGTGGATTACGGAGATTAATCT 697  
DB 287 ATGCGCTAAGTCTCAAGCAGCCGAACTTTGAGACTCATTTAAGCAGGGGAGCTT 346  
QY 698 ACTTCTCTTCAGGGAATAGCAGTGGAGTATAACACCATGGAAGAGTATTTCCAA 757  
DB 347 ATTTCTCTTTTCGGAAACCGCGTGTGATGATCACTGTCGCAAGCGGATTTAT 405  
QY 758 GAGTGGCTCAGTTTGTAAAGATGATATGGAGAGATCTCAAGAGATGCTGAGAAAG 817  
DB 407 GCGTTGCCCGGCTGTCGAAATGGGACAAAGGTGGCGGCGATCGATTC---GCA 463  
QY 818 GGAGCTGCTTCTGAAGGCGGCTGAACTGCTCAGTTCCTGGAGAGATCTCAITTTAT 877  
DB 464 GCACATCTTCTCAAGGCGGCTCAACAGCTCCATTCGCGGAGATATCTTCTACT 523  
QY 878 TCACATCTCTCAAGGAGTACAGATGAT-----TCGTATCAAGGCGGTG 925  
DB 524 TTAATGAATCCAAATGCGCAAGATGCGGAGGAGGAGTATGATGCAAGAGTCA 583  
QY 926 ATGTGTGCTGGCAAGCTTCTACACCTTAACAGGAGGAGGAGTATGATGCAAGAG 985  
DB 584 AACTGATCTACGAGTCTTCAACAGGCGGAGCAACTCAATCCGCGCTCAGGCTTTC 643  
QY 986 CCAATGATGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 1045  
DB 644 CTTTGGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAGGAG 703  
QY 1046 CCAATGATGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 1105  
DB 704 TCAACTGCAAGTCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAG 763  
QY 1106 GCGTGGCTGCAAGGAGTCTTCAACAGGAGGAGGAGTATGATGCAAGAGGAG 1165  
DB 764 GTCACAGGATTCG-----AGAGGCTTCCGGATCCGAGCAG 799





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:44:33 : Search time 12072.7 Seconds  
(without alignments)  
11427.132 Million cell updates/sec

Title: us-09-856-681-1  
Perfect score: 3093  
Sequence: 1 atyagtgcaagagcctt.....acaaatagcgtgacataa 3093

Scoring table: IDENTITY\_NUC  
Gapop 10.0 / Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 577422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 6%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:
- 2: gb\_ba:
- 3: gb\_ba:
- 4: gb\_ba:
- 5: gb\_ov:
- 6: gb\_pat:
- 7: gb\_pat:
- 8: gb\_pl:
- 9: gb\_pl:
- 10: gb\_ro:
- 11: gb\_ro:
- 12: gb\_sy:
- 13: gb\_uni:
- 14: gb\_uni:
- 15: gb\_uni:
- 16: gb\_uni:
- 17: gb\_uni:
- 18: gb\_uni:
- 19: gb\_uni:
- 20: gb\_uni:
- 21: gb\_uni:
- 22: gb\_uni:
- 23: gb\_uni:
- 24: gb\_uni:
- 25: gb\_uni:
- 26: gb\_uni:
- 27: gb\_uni:
- 28: gb\_uni:
- 29: gb\_uni:
- 30: gb\_uni:
- 31: gb\_uni:
- 32: gb\_uni:
- 33: gb\_uni:
- 34: gb\_uni:
- 35: gb\_uni:
- 36: gb\_uni:
- 37: gb\_uni:
- 38: gb\_uni:
- 39: gb\_uni:
- 40: gb\_uni:
- 41: gb\_uni:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3093	100.0	3093	6	AX026741	Sequence
2	3093	100.0	3862	6	AX026746	Sequence
3	3093	100.0	3862	9	AF279656	Homo sapi
4	3032	98.0	4250	9	AB037789	Homo sapi
5	2929.6	96.0	6080	6	BD160721	Primer fo
6	2929.6	96.0	6080	9	AK027867	Homo sapi
7	2746.2	88.8	3550	6	AX095520	Sequence
8	2414.8	78.1	3018	10	AF288666	Mus muscu
9	2182.4	70.6	3041	6	BD159617	Primer fo
10	2182.4	70.6	3041	9	AK027501	Homo sapi
11	2137.4	65.1	2770	10	AF030430	Mus muscu
12	1824.4	59.0	2293	6	BD157826	Primer fo
13	1824.4	59.0	2293	9	AK027471	Homo sapi
14	1742.4	56.3	2227	6	BD127394	Primer fo
15	1742.4	56.3	2227	9	AK074975	Homo sapi
16	1666.4	53.9	2306	6	BD155853	Primer fo
17	1666.4	53.9	2306	9	AK027654	Homo sapi
18	1462	47.3	2123	6	BD155987	Primer fo
19	1462	47.3	2123	9	AK027439	Homo sapi
20	1367.2	44.2	3634	9	AK096337	Homo sapi
21	1200	38.6	13439	2	AC010497	Homo sapi
22	1200	38.6	13439	9	AC010296	Homo sapi
23	1200	38.6	134061	9	AC008524	Homo sapi
24	1200	38.6	179647	2	AC108124	Homo sapi
25	1200	38.6	188207	2	AC102333	Homo sapi
26	1023.8	33.1	170385	2	AC124181	Mus muscu
27	1023.8	33.1	184365	10	AC121783	Mus muscu
28	1023.8	33.1	237974	2	AC094771	Rattus no
29	1023.8	33.1	253924	2	AC126707	Rattus no
30	1023.8	33.1	256786	2	AC125773	Rattus no
31	788	25.5	833	6	BD147394	Primer fo
32	758.6	24.5	814	6	BD125056	Primer fo
33	758.6	24.5	814	6	BD127012	Primer fo
34	751.2	24.3	780	6	BD147572	Primer fo
35	742.6	24.0	788	6	BD126945	Primer fo
36	681.8	22.0	6161	10	AK122515	Mus muscu
37	672.2	21.7	5886	9	AF389427	Homo sapi
38	672.2	21.7	5943	9	AF389428	Homo sapi
39	672.2	21.7	5111	9	AF389429	Homo sapi
40	669	21.6	4234	6	AX492975	Sequence
41	661.6	21.4	3205	6	AX685967	Sequence
42	659	21.3	5924	9	AB040912	Homo sapi
43	659	21.3	5925	9	AF389430	Homo sapi
44	657.4	21.3	662	6	BD150362	Primer fo
45	642	20.8	2191	6	AX685969	Sequence

ALIGNMENTS

AX026741 AX026741 4093 bp DNA linear PAT 16-SEP-2000  
Sequence : from Patent WO0031252.  
AX026741  
AX026741.1 GI:10147886  
Homo sapiens (Human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Klostermann A. and Wehl C.  
Human semaphorin 6a-1 (sema6a-a), a gene involved in neuronal  
development and regeneration mechanisms during apoptosis, and its

use as a potential drug target  
 Patent: WO 0031252-A 1 02-JUN-2000  
 KLOSTERMANN ANDREAS (DE) ; MAX PLANCK ZELLENS-BANK (DE) ; BEH.  
 CHRISTIAN (DE)

# FEATURES

Location/Qualifiers

1..3093  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

1..3093  
 /note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC03107.1"

/db\_xref="GI:10157887"

/translation="MRPALLLYTLHFAGAGPEPDEPSEPTISISGHNVTROYFVFGH  
 KPRNTTORHRLDIOMIMNGTLYIAARDHYIVTDIDTSHIEIYCSKLLIMKSKQA  
 DVTCDRMKGKHKHCHNEPIKLLKNDDELEVCNINAPNPNCKMOTLEPTGDDFS  
 GMARCPYDAKHANVALFADGKLYSATVTDALDAIVRSIGESPTLRTVKHDSKWLK  
 RPFVQAVDYGDIYVFFREIAVENTMCKVPEPVAQVCKNDGSSORVLEKQWLF  
 LKARLNGSPVDSHFEYENILQAVTDVIRINGROVLATEFEPYNSIPGSAYCAIMLD  
 IASVTGFKKQKSPDS:WTPVPERVPRPGCCAGSSSLERATSEPTIDPTLRFI  
 KTHPLMDKAVPSIFRPMFLKIMVRYRLTKIADVTAAGPYQNHVTVPLGSEKGIILKF  
 LARICNSGFUNDLFLREMSYINSEKSYDGVDRKRMKGLDRASSLYVAFSTCV  
 KYPLECRCHRGSKKCTASNDPYGWIKFEGAGSHLSPNSRLTFEGDIERGNTDSIG  
 DCHNSFVALHSSSELPSFTTSDSTAGEGVESRGMDKMLHLLDPSDIPLGAVSS  
 HHQDKGVIRSESLYKHDQVLPVTLALAVLAFVMGAFSGIYVVCVCHRRKQVA  
 VVORKEKELTHSRSGSSSVTKISGLFCTQSKOPKPEALILPEMHNKKILATPQNAK  
 MLIKADQHHLLTALPTPESTFTLOKKRKSREKERNKLNILNACTKMPMTGSPV  
 ITIDPLRSPSHIYVVVILPITOOGYOHVEYVDFKMSSEVAOMALEDDAATI EYKTK  
 EHLSSKSNHGVN:VENIDSLFPPVQREASLGFPGASLSQGTLSKRLXKHHSSTGV  
 DYKRSYPSNLTTRHOATIKPNNINSSNSHLSRNSFGGDKPPAPQRFVSLQVR  
 SQQPSQCAVTVSRQPSLNAYNSLTRSLKRLFTSLKPDVPPPKPSEAPLSTSMKPNDACT

BASE COUNT 813 a 855 c 765 g 660 t

## ORIGIN

Query Match 100.0% Score 3093; DR 6; Length 3093;

Best Local Similarity 100.0% Pred. No. 0;

Matches 3093; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTCAGAGCTTCTGCTGATATTCACACAGTGGIACACTTTCGUGGGTGGTTC 60  
 DB 1 ATGAGCTCAGAGCTTCTGCTGATATTCACACAGTGGIACACTTTCGUGGGTGGTTC 60  
 QY 61 CCAGAGATTTCTGAGCCATCAGTATTCGCAATGCAATATACAAAACATATTCAGTGS 120  
 DB 61 CCAGAGATTTCTGAGCCATCAGTATTCGCAATGCAATATACAAAACATATTCAGTGS 120  
 QY 121 TTGTGGGCGCACAGCCAGAGCGGACACACACAGAGGACAGATTCAGATTCAGAT 180  
 DB 121 TTGTGGGCGCACAGCGAGGACACACACAGAGGACAGATTCAGATTCAGATTCAGAT 180  
 QY 181 ATTATGATCAAGCGAGACGCTACATTCGCTGATGAGACATATATATATATATAT 240  
 DB 181 ATTATGATCAAGCGAGACGCTACATTCGCTGATGAGACATATATATATATATATAT 240  
 QY 241 ATGACACATCAGCAGAGAGATTTATTTACGAAACATAGCAGTGAATTCAGAT 300  
 DB 241 ATGACACATCAGCAGAGAGAGATTTATTTGACAAAGAAATGACATGGAATTCAGAT 300  
 QY 301 CAGGCGGATGTACACATTCAGATTCAGAGGAGAAACATAGAGATGAGTGCATACAT 360  
 DB 301 CAGGCGGATGTACACATTCAGAGATTCAGAGGAGAAACATAGAGATGAGTGCATACAT 360  
 QY 361 ATTAAGATCTTCTTAAGAAACATGATGATGATTTGCTGTGAGAACTAAAGCCTTC 420  
 DB 361 ATTAAGATCTTCTTAAGAAACATGATGATGATTTGCTGTGAGAACTAAAGCCTTC 420  
 QY 421 AACCTTCTCTGACAGAACTATAGATGATGATGATTTGAGAACCATCTCGGAGATTC 480  
 DB 421 AACCTTCTCTGACAGAACTATAGATGATGATGATTTGAGAACCATCTCGGAGATTC 480  
 QY 481 GGAATGCCAGATGCGGCATATGATGCCAAACATGCCAGCTGTCACATGTTTTCAGATGGA 540

DB 481 GGAATGCCAGATGCGGCATATGATGCCAAACATGCCAGCTGTCACATGTTTTCAGATGGA 540  
 QY 541 AATATATATCTGAGCAGAGTCTACTATCTCTTCCGATTTACCGAGTCATTTACCGAGT 600  
 DB 541 AATATATATCTGAGCAGAGTCTACTATCTCTTCCGATTTACCGAGTCATTTACCGAGT 600  
 QY 601 CTGGAGAAAGCTCTACTCTGAGGACCTGCAAGCAGGATTCAAANTGTTGAAGAACCA 660  
 DB 601 CTGGAGAAAGCTCTACTCTGAGGACCTGCAAGCAGGATTCAAANTGTTGAAGAACCA 660  
 QY 661 TACTTGTTCAGGCGTGSATTCAGGAGATTAATCTACTTCTTCTTCCAGGGAATAGCA 720  
 DB 661 TACTTGTTCAGGCGTGSATTCAGGAGATTAATCTACTTCTTCTTCCAGGGAATAGCA 720  
 QY 721 GTGGAGATTAACACCAICGGAAAGTACTTTTCCAGAGTGGCTCAGGTTTGTGAAGAT 780  
 DB 721 GTGGAGATTAACACCAICGGAAAGTACTTTTCCAGAGTGGCTCAGGTTTGTGAAGAT 780  
 QY 781 GATATGGGAGGATCTCAAGAGTCTCTGGAGAAACAGTGGAGTGGTCTCTTCCAGAGGCGGC 840  
 DB 781 GATATGGGAGGATCTCAAGAGTCTCTGGAGAAACAGTGGAGTGGTCTCTTCCAGAGGCGGC 840  
 QY 841 TTGAATCTGCTCAGTTCCTTGGAGACTCTCATTTTATTTTCAACATCTCCAGGCGAGTAC 900  
 DB 841 TTGAATCTGCTCAGTTCCTTGGAGACTCTCATTTTATTTTCAACATCTCCAGGCGAGTAC 900  
 QY 901 GATGATCTGATCAAGGGGGGAGTGTCTCTGGCAAGCTTCTTACACCTTATTAAC 960  
 DB 901 GATGATCTGATCAAGGGGGGAGTGTCTCTGGCAAGCTTCTTACACCTTATTAAC 960  
 QY 961 AGCATCCCTGGTCTGCACTCTGTCCTATGACATGCTGACATGCTGACATGCTGACAT 1020  
 DB 961 AGCATCCCTGGTCTGCACTCTGTCCTATGACATGCTGACATGCTGACATGCTGACAT 1020  
 QY 1021 GGGAGATTCAGAGAAACAGAGTCTCTGATTCACCTCCAGCAGGATTCCTGATGAACGA 1080  
 DB 1021 GGGAGATTCAGAGAAACAGAGTCTCTGATTCACCTCCAGCAGGATTCCTGATGAACGA 1080  
 QY 1081 GTCTCTAAGCCAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 DB 1081 GTCTCTAAGCCAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 QY 1141 TCCATGATTTCCCTGATGA:ACCCGAGACTTCATCAAGAGCAGCAGCAGCAGCAGCAG 1200  
 DB 1141 TCCATGATTTCCCTGATGATACCCGAGACTTCATCAAGAGCAGCAGCAGCAGCAGCAG 1200  
 QY 1201 GGNATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 DB 1201 GGNATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 1261 ACCAAATTCAGTGSAGACAGCTGCTGGGCGCATATCAGATCAGACTGCTGCTGCTGCTGCT 1320  
 DB 1261 ACCAAATTCAGTGSAGACAGCTGCTGGGCGCATATCAGATCAGACTGCTGCTGCTGCTGCT 1320  
 QY 1321 GATCAGAGAGGAGATCATCTGAGGTTTTTGGCAGANTAGGANTAGGTTTTCTA 1380  
 DB 1321 GATCAGAGAGGAGATCATCTGAGGTTTTTGGCAGANTAGGANTAGGTTTTCTA 1380  
 QY 1381 AATGACAGCTTTTCTGAGGAGATGAGTGTTCAGACTCTGAAATGCAGCTATGAT 1440  
 DB 1381 AATGACAGCTTTTCTGAGGAGATGAGTGTTCAGACTCTGAAATGCAGCTATGAT 1440  
 QY 1441 GGAGTGGAGACAAAGAGTATCGGCATGCGAGTGGACAGAGCAAGCAGCTCTCTGAT 1500  
 DB 1441 GGAGTGGAGACAAAGAGTATCGGCATGCGAGTGGACAGAGCAAGCAGCTCTCTGAT 1500  
 QY 1501 GTTGCT 1560  
 DB 1501 GTTGCT 1560  
 QY 1561 TGTAAAAAAGCTGTATTGCTCTCCAGAGACCATATTTGATGATGATAAAGAGAGTGGT 1620  
 DB 1561 TGTAAAAAAGCTGTATTGCTCTCCAGAGACCATATTTGATGATGATAAAGAGAGTGGT 1620



Query Match									
Best Local Similarity 100.0%; Pred. No. G:									
Matches 3093; Conservative C: Mismatches G: Indels G: Gaps G:									
QY	1	ATGAGTCAGAGCCCTTGCTGCTA-TATTCACACTGGTACACTTGTGCTGGGCTGGTTC	63						
DB	658	ATGAGTCAGAGCCCTTGCTGCTA-TATTCACACTGGTACACTTGTGCTGGGCTGGTTC	717						
QY	61	CCAGAAGATCTTACGCAATCAGTATTTGGCATGCGCAACTATACAAACACATATCCGCTG	120						
DB	718	CCAGAAGATCTTACGCAATCAGTATTTGGCATGCGCAACTATACAAACACATATCCGCTG	777						
QY	121	TTTGTSGCCACAGCCAGGAGCGGACACACACAGAGCGACGCTGGACATCCAGATG	280						
DB	778	TTTGTSGCCACAGCCAGGAGCGGACACACACAGAGCGACGCTGGACATCCAGATG	837						
QY	181	ATTATGATCATGAACGGAACCTCTACATTTGTGCTAGGACCATATTATTTACTGTTCAT	240						
DB	838	ATTATGATCATGAACGGAACCTCTACATTTGTGCTAGGACCATATTATTTACTGTTCAT	897						
QY	241	ATAGACACATTCACACACGAGAGAMTTTATTTGCAAAAACACTTACATTTGAATCTAGA	300						
DB	898	ATAGACACATTCACACACGAGAGAAATTTATTTAGCAAAAACACTTACATTTGAATCTAGA	957						
QY	301	CAGGCGGATGTACACACATCAGATTAAGTGAAGGAAACATTAAGGATGCTGCCAAGCTTT	360						
DB	958	CAGGCGGATGTACACACATCAGATTAAGTGAAGGAAACATTAAGGATGCTGCCAAGCTTT	1017						
QY	361	ATTAAGTCTCTTAAAGAAACGATGATGCTTTGCTGTGGACATTAAGCTTC	420						
DB	1018	ATTAAGTCTCTTAAAGAAACGATGATGCTTTGCTGTGGACATTAAGCTTC	1077						
QY	421	AACCTTCTGCGAAGATATTAAGTGGATACATTTGGAACCATTCGGGATCAATTCAGC	480						
DB	1078	AACCTTCTGCGAAGATATTAAGTGGATACATTTGGAACCATTCGGGATCAATTCAGC	1137						
QY	481	GGAATGGCAGATGGCCATATGATGCAAAACATGGAACGTTGACATTTTCAGATGGA	540						
DB	1138	GGAATGGCAGATGGCCATATGATGCAAAACATGGAACGTTGACATTTTCAGATGGA	1197						
QY	541	AAACTATATCAGCAGAGTACATTCCTTTCATATTCAGGATCTCATTTACCGGAT	600						
DB	1198	AAACTATATCAGCAGAGTACATTCCTTTCATATTCAGGATCTCATTTACCGGAT	1257						
QY	601	CTTGAGAAAGCCCTACGCTGGGAGCGGCAAGCAGATTCAAAATGTTGAAAGAAACA	660						
DB	1258	CTTGAGAAAGCCCTACGCTGGGAGCGGCAAGCAGATTCAAAATGTTGAAAGAAACA	1317						
QY	661	TACTTTGTTCAAGCCCTGGATACGAGATATATCTATCTTTTCAGGAAATAGCA	720						
DB	1318	TACTTTGTTCAAGCCCTGGATACGAGATATATCTATCTTTTCAGGAAATAGCA	1377						
QY	721	GTGAGATATTAACACCATGGGAAGGTAGTTTCCCAAGATGGGTTCAGCTTTGAAAGAT	780						
DB	1378	GTGAGATATTAACACCATGGGAAGGTAGTTTCCCAAGATGGGTTCAGCTTTGAAAGAT	1437						
QY	781	GATATGGGAGATCTCAAGAGTCTCTGGAGAAACAGTGGAGCTGCTTCAAGAGCGGC	840						
DB	1438	GATATGGGAGATCTCAAGAGTCTCTGGAGAAACAGTGGAGCTGCTTCAAGAGCGGC	1497						
QY	841	ITCAAGTGTCTAGTCTCTGAGTCTCTATTTTATTTCAACATTTCCAGCGAGTACA	900						
DB	1498	ITCAAGTGTCTAGTCTCTGAGTCTCTATTTTATTTCAACATTTCCAGCGAGTACA	1557						
QY	901	GATGCTGATTATCAACCGCGGTGATGTTGTCTCGGCAACGTTTCTACACCTTATTAAC	960						

DB	1558	GATGCTGATTATCAACGGCGGTGATGTTGCTGGCAAGCTTTTCTACACCTTATAAC	1617						
QY	961	AGCATCCCTGGCTTGGAGCTGTGCTGATGATGATGCTTGGAGCTTGGAGCTTGGAGCT	1620						
DB	1628	AGCATCCCTGGCTTGGAGCTGTGCTGATGATGATGCTTGGAGCTTGGAGCTTGGAGCT	1677						
QY	1021	GGAGATTCAGAGAAAGATGCTGCTGATTCACCTGGAGACCAAGTTCCTGATGAACGA	1080						
DB	1678	GGAGATTCAGAGAAAGATGCTGCTGATTCACCTGGAGACCAAGTTCCTGATGAACGA	1737						
QY	1031	GTTCCTAAGCCAGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140						
DB	1738	GTTCCTAAGCCAGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1797						
QY	1141	TCCATAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1200						
DB	1798	TCCATAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1857						
QY	1201	GGAGTGGCTGCTGATTCACAGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1260						
DB	1858	GGAGTGGCTGCTGATTCACAGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1917						
QY	1261	ACCAAAATTCAGTGGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320						
DB	1918	ACCAAAATTCAGTGGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1977						
QY	1321	GGATCAGAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1380						
DB	1978	GGATCAGAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2037						
QY	1381	AATGACAGCTTTTCTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT	1440						
DB	2038	AATGACAGCTTTTCTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT	2097						
QY	1441	GGAGTGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1500						
DB	2098	GGAGTGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2157						
QY	1501	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1560						
DB	2158	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2217						
QY	1561	TGTAAAGAAACCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620						
DB	2218	TGTAAAGAAACCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2277						
QY	1621	GGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1680						
DB	2278	GGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2337						
QY	1681	AATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1740						
DB	2338	AATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2397						
QY	1741	TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800						
DB	2398	TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2457						
QY	1801	GGAGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1860						
DB	2458	GGAGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2517						
QY	1861	GGAGTGTCTTCCATATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1920						
DB	2518	GGAGTGTCTTCCATATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2577						
QY	1921	GGCCAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980						
DB	2578	GGCCAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2637						
QY	1981	ATGGGGGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040						



Db	718	CCAGAGATCTCTGACCAATACAGTATTCTTGCAATGGCAATATATACAAATAGATTGCGGCTG	777
Qy	121	TTTGTGGGACAAAGCCAGTACGGAACACACACACAGAGACATAGGTTGGACATCCACAGATG	120
Db	778	TTTGTGGGACAAAGCCAGTACGGAACACACACACAGAGAGGCTGGACATCCACAGATG	837
Qy	181	ATTATGATCATGAAAGGAAAGCTCTACATTCGCTGCTAGGACACCAATTTATATAGCTGTGAT	249
Db	838	ATTATGATCATGAAAGGAAAGCTCTACATTCGCTGCTAGGACACCAATTTATATAGCTGTGAT	897
Qy	241	ATAGACATCTACACACAGGAGAAATTTATTGTACCAAAAAACTTGACATGGAAATCTAGA	300
Db	898	ATAGACATCTACACACAGGAGAAATTTATTGTACCAAAAAACTTGACATGGAAATCTAGA	957
Qy	301	CAGGCCGATGTAGACACATTCACAGATGAAGGAAACATAAAGGATGAGTGCACACACTT	360
Db	958	CAGGCCGATGTAGACACATTCACAGATGAAGGAAACATAAAGGATGAGTGCACACACTT	1017
Qy	361	ATTAAAGTCTTCTAAAGAAACAGGATGGCATTTTCTCTGTTGGAACCTATATGCTTTC	420
Db	1018	ATTAAAGTCTTCTAAAGAAACAGGATGGCATTTTCTCTGTTGGAACCTATATGCTTTC	1077
Qy	421	AACCTTCTCTGCAGAAACTATAGATGGATACATATGGAACCATTCGGGGAATGAAATTCAGC	480
Db	1078	AACCTTCTCTGCAGAAACTATAGATGGATACATATGGAACCATTCGGGGAATGAAATTCAGC	1137
Qy	481	GGAAATGCCAGATGCCCATATGATGCCAAACATGCCAATCTTCAGACTCTTCCAGATGGA	540
Db	1138	GGAAATGCCAGATGCCCATATGATGCCAAACATGCCAATCTTCAGACTCTTCCAGATGGA	1197
Qy	541	AAACTATATACGCCACAGTACTGACTTCTCTCCCATTTGACCCAGTCAATTTACCGGAGAT	600
Db	1198	AACTATATACGCCACAGTACTGACTTCTCTCCCATTTGACCCAGTCAATTTACCGGAGAT	1257
Qy	601	CTTGAGAAAGCCCTACCCCTCGGACCGTCAAGACGATTCAAAATGCTTGAAGAAACCA	660
Db	1258	CTTGAGAAAGCCCTACCCCTCGGACCGTCAAGACGATTCAAAATGCTTGAAGAAACCA	1317
Qy	661	TACTTTGTTCAAGCCGCTGGATTACGGAGATTATCTACTTCTTCTTCAGGGAATTAAGCA	720
Db	1318	TACTTTGTTCAAGCCGCTGGATTACGGAGATTATCTACTTCTTCTTCAGGGAATTAAGCA	1377
Qy	721	GTGGAGTATACACCATTCGGAAAGGTACTTTCCGACAGATGCTCAGGTTTCTAAGAAAT	780
Db	1378	GTGGAGTATACACCATTCGGAAAGGTACTTTCCGACAGATGCTCAGGTTTCTAAGAAAT	1437
Qy	781	GATAIGGAGGATCTCAAGAGTCTCGAGAAACATGGAGTCGCTTCTCTGAAGCGGCGC	840
Db	1438	GATAIGGAGGATCTCAAGAGTCTCGAGAAACATGGAGTCGCTTCTCTGAAGCGGCGC	1497
Qy	841	TTGAATCTGCTCAGTCTCTGAGAGACTCTGTTTTTATTCACAACTCTCCAGGCGAGTTACA	900
Db	1498	TTGAATCTGCTCAGTCTCTGAGAGACTCTGTTTTTATTCACAACTCTCCAGGCGAGTTACA	1557
Qy	901	GATGAGATTGGTATCAACGGGCGTATGTGTCTGCAACAGTTTCTACAGCTATTAAG	960
Db	1558	GATGAGATTGGTATCAACGGGCGTATGTGTCTGCAACAGTTTCTACAGCTATTAAG	1617
Qy	961	AGCATCCCTGGGCTGCGAGTCTGTGCGATATGACATGCTTGACATTCGCGATGTTTTAAT	1020
Db	1618	AGCATCCCTGGGCTGCGAGTCTGTGCGATATGACATGCTTGACATTCGCGATGTTTTAAT	1677
Qy	1021	CGGAGATTCAGGAGACAGATCTCTGATTCGACCTGAAACAGATTCCTGTGATGACCA	1080
Db	1678	CGGAGATTCAGGAGACAGATCTCTGATTCGACCTGAAACAGATTCCTGTGATGACCA	1737
Qy	1081	GTTCCTTAAGGCCAGGCCAGGTTGCTGTGCTGGCTTCACTCTCTCTAGAAAGATATGCAACC	1140
Db	1738	GTTCCTTAAGGCCAGGCCAGGTTGCTGTGCTGGCTTCACTCTCTCTAGAAAGATATGCAACC	1797
Qy	1141	TCCAAATGAGTTCCTGTGATGATACCTGAACTTATCAAGAGCGACCCGCTCATGGAATGAG	1200
Db	1798	TCCAAATGAGTTCCTGTGATGATACCTGAACTTATCAAGAGCGACCCGCTCATGGAATGAG	1857

[illegible]











```

RESUL1 6
AK027867 6360 bp mRNA linear PR: 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ14961 fis, clone PACE100230, highly similar
DEFINITION tc Mus musculus semaphorin 7a mRNA.
ACCESSION AK027867 1 GI:14042853
VERSION cDNA capping, fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1
AUTHORS Isozaki,T., Ota,T., Hayashi,K., Suiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sudano,S., Takahashi,Fujii,A., Hara,H.,
Tanabe,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nakahara,K., Masuko,Y. and Oshima,A.
NEDO human cDNA sequencing project.
TITLE Unpublished
REFERENCE 2 (bases 1 to 6360)
AUTHORS Isozaki,T. and Otsuki,T.
JOURNAL Direct Submission
TITLE Submitted (10-MAY-2003) Takao Isozaki, Helix Research Institute,
Genomics Laboratory, 1542-3 Yama, Kisarazu, Chiba 292-0812, Japan
COMMENT (E-mail:genomics@hri.co.jp, Tel:01-438-52-9975, Fax:01-438-52-3586)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full library construction;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
location/Qualifiers
.. 6060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="P:ACE100230"
/sequence_type="placenta"
/clone_lib="v1ACE4"
/notes="Cloning vector: pMEL3SF12"
89..1345
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB55418.1"
/db_xref="GI:14042853"
/translating="MIMNGUYIAAROHYIVDID7SHTEIYCSKKLTWKSQAQVD
TCRMKGRKDECNFIKVLKKNDALFVCGTNFNPSCRNKMDTLEPPGDFSGMA
RCPYDAKHANVALFADGKLYSATVDFLAIDAVIYRSLGESPTLRIVKDSKWLKEPY
FYQAVDGYDYIEFRELAVEYNTMGKVVFPRVAQVCKNDGGSORVLEKQWTSFLKA
RNCQSVPGDSHFYFNLCGAVTDV: RINGRDVLA7FSTPNSIPGSACVADMLDIAS
VLTGRFKQKSPES:WTPVDERVPKPRGCCAGSSSLERATSNRFPDDILNFIKTH
PLMDEAVPS:INRPFIL:MYRYRLIK:AVDIAGPYONHIVFVDSKGL:LLFLAR
IGNSQEMDSFLEPMSVINSFKMQL"
BASE COUNT 1639 a 1526 c 1363 g 1532 t
ORIGIN
Query Match 96.0%; Score 2569.6; DB 9; Length 6060;
Best local Similarity 99.8%; Pred. No. 0;
Matches 2994; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 96 CAACTATACAAACAGTATCGCGTGTCTTGTGGCCACAGCCAGCGACGGAACACACACA 155
DB 1 CAACTATACAAACAGTATCGCGTGTCTTGTGGCCACAGCCAGCGACGGAACACACACA 60
QY 156 GAGGCACAGGCTGGACATCCAGATGATTATGATCATGATGACGACGACCCCTCTACATGCTGCC 215
DB 61 GAGGCACAGGCTGGACATCCAGATGATTATGATCATGATGACGACGACCCCTCTACATGCTGCC 120
QY 216 TAGGCACCATATTTACTGTTGATATGACATCATGACACGAGGAAGAAATTTATGTAG 275
DB 1201 TCAGAAATCACACTGTGGTCTTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTTTGGC 1260
DB 121 TAGGACCATATTTATACIGTGTGATATAGACACAACACACACGGAAGAAATTTATGTAG 180
QY 276 CAAAAAATGACATGGAAATCTAGACAGGCGGATGTAGACACATGCGAAATGAAGGAAA 335
DB 181 CAAAAAATGACATGGAAATCTAGACAGGCGGATGTAGACACATGCGAAATGAAGGAAA 240
QY 336 ACATAAGGATGAGTCCACACACATTTATTAAGTTCTTCT:AAAGAAACACATGATGATT 395
DB 241 ACATAAGGATGAGTCCACACACATTTATTAAGTTCTTCT:AAAGAAACACATGATGATT 300
QY 396 GTTGTGTGTGGAACTAATGCCITCAACCCCTTCTGAGAAACTATAGATGGATACATT 455
DB 301 GTTGTGTGTGGAACTAATGCCITCAACCCCTTCTGAGAAACTATAGATGGATACATT 360
QY 455 GGACACATTCGGGATGAATTCAGCGGAATGGCCAGATGCGCATATGATGCCAAACATGC 515
DB 361 GGACACATTCGGGATGAATTCAGCGGAATGGCCAGATGCGCATATGATGCCAAACATGC 420
QY 516 CAACGTTGCACCTGTTGCAGATGGAAACTATATCTCAGCCACAGTGAATCTCCTTCC 575
DB 421 CAACGTTGCACCTGTTGCAGATGGAAACTATATCTCAGCCACAGTGAATCTCCTTCC 480
QY 576 CATTCAGCACTCATTTAGGAGAGTCTTGGAGAAAGCCCTACCCCTGGGACCGTCAAGCA 635
DB 481 CATTCAGCACTCATTTAGGAGAGTCTTGGAGAAAGCCCTACCCCTGGGACCGTCAAGCA 540
QY 636 CGATTCAAAATGTTGAAGAACCATACTTGTTCAGACCGTGGATTCAGGAGATATAT 695
DB 541 CGATTCAAAATGTTGAAGAACCATACTTGTTCAGACCGTGGATTCAGGAGATATAT 600
QY 696 CTACTTCTCTCAGGAAATACAGTGGAGTATAATACCATGGAAAGTAGTTTTCCT 755
DB 601 CTACTTCTCTCAGGAAATACAGTGGAGTATAATACCATGGAAAGTAGTTTTCCT 650
QY 756 AAGAGTGGCAGCTTGTAGAAATATATGGAGAGTCTCAAGAGTCTCTGAGAAACA 815
DB 661 AAGAGTGGCAGCTTGTAGAAATATATGGAGAGTCTCAAGAGTCTCTGAGAAACA 720
QY 816 GTGAGCTGCTTCGAGGCGCTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
DB 721 GTGAGCTGCTTCGAGGCGCTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 876 TTCAACATCTCTCAGCAGTATACAGATGTGATCTGATCAAGGGGGGTGATGTGCTCT 935
DB 781 TTCAACATCTCTCAGCAGTATACAGATGTGATCTGATCAAGGGGGGTGATGTGCTCT 840
QY 936 GGCAACCTTTCTACACCTTATACAGATCCCTGGGTCTGAGTCTGCTGCTGCTGCTGCT 995
DB 841 GGCAACCTTTCTACACCTTATACAGATCCCTGGGTCTGAGTCTGCTGCTGCTGCTGCT 900
QY 996 GCTTGACATTCGCGAGTGTCTTCTGCGAGATTCAGAGAACAGAGTCTCTGATTCAC 1055
DB 901 GCTTGACATTCGCGAGTGTCTTCTGCGAGATTCAGAGAACAGAGTCTCTGATTCAC 960
QY 1056 CTGACACACAGTCTCTGATGAAGAGTGTCTTAAGCCCGAGCCAGTGTGCTGCTGCTC 1115
DB 961 CTGACACACAGTCTCTGATGAAGAGTGTCTTAAGCCCGAGCCAGTGTGCTGCTGCTC 1020
QY 1116 ATCTCTCTTAGAAGATATGACCTCAATGATGATGATGATGATGATGATGATGATGAT 1175
DB 1021 ATCTCTCTTAGAAGATATGACCTCAATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1176 CAAGACACCGCTCATGATGAGGAGTGGCTCTCATCTTCAACAGGCGATGGTTCCT 1235
DB 1081 CAAGACACCGCTCATGATGAGGAGTGGCTCTCATCTTCAACAGGCGATGGTTCCT 1140
QY 1236 GAGAACAAATGTCATACCGCTTACCAAAATGTCAGTGGACACAGTCTCTGGGCGATA 1295
DB 1141 GAGAACAAATGTCATACCGCTTACCAAAATGTCAGTGGACACAGTCTCTGGGCGATA 1200
QY 1296 TCAGAAATCACACTGTGGTCTTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTTTGGC 1355
DB 1201 TCAGAAATCACACTGTGGTCTTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTTTGGC 1260

```



QY 1 ATCAJGTCAAGACCGTTGCGATATATTCACACTGCTACACTTTGTGGGGCTGGTTC 60  
DB 126 ATGAGTTCAGAAAGCTTGGCTATATATTCACACTGCTACACTTTGTGGGGCTGGTTC 185  
QY 61 CCAGAAAGATTCTGACCCAAATGAGTATTCGCAATGCAATATACAAAACACTATCCGCTG 120  
DB 156 CCAGAAAGATTCTGACCCAAATGAGTATTCGCAATGCAATATACAAAACACTATCCGCTG 245  
QY 121 TTGTGGGCGACAGACCGAGCGGACACACACAGACAGACAGACAGACAGACAGACAG 180  
DB 246 TTGTGGGCGACAGACCGAGCGGACACACACAGACAGACAGACAGACAGACAGACAG 305  
QY 181 ATTATGATCATCAAGGGAAACCTCTACATTTGCTGTAGGACCATATTTACTCTGAT 240  
DB 306 ATTATGATCATCAAGGGAAACCTCTACATTTGCTGTAGGACCATATTTACTCTGAT 365  
QY 241 ATAGACATCATCACACGAGAGAAATTTATTTAGCAAAANAACTGCACATGNAATCTATA 300  
DB 366 ATAGACATCATCACACGAGAGAAATTTATTTAGCAAAANAACTGCACATGNAATCTATA 425  
QY 301 CAGGCGGATGTAGACATCCAGATGAAGGGAAACATAAAGSNTAGTGGACACACTT 180  
DB 426 CAGGCGGATGTAGACATCCAGATGAAGGGAAACATAAAGSNTAGTGGACACACTT 485  
QY 361 ATTAAGATTCTCTAAAGAAACAGATCATCATTTGCTGTAGGACCATATTTACTCTGAT 420  
DB 486 ATTAAGATTCTCTAAAGAAACAGATCATCATTTGCTGTAGGACCATATTTACTCTGAT 545  
QY 421 AGCCTTCTGTGAGAAACTATAGATGATAGATGATAGATGATGATGATGATGATGATG 480  
DB 546 AGCCTTCTGTGAGAAACTATAGATGATAGATGATAGATGATGATGATGATGATGATG 605  
QY 481 GGAATGGGACAGTGGCATATGATGCCAAACATGCGAAGCTGGACTTTTGCATATGGA 540  
DB 606 GGAATGGGACAGTGGCATATGATGCCAAACATGCGAAGCTGGACTTTTGCATATGGA 565  
QY 541 AAATATAGTCAAGCACAGTCACTGACTTCCTTGGCATTTCAGCAGACATTTACCGGAGT 600  
DB 566 AAATATAGTCAAGCACAGTCACTGACTTCCTTGGCATTTCAGCAGACATTTACCGGAGT 725  
QY 601 CTGGAGAAAGCGCTACCGCTGGGACCGTCAGACAGATTCAAAATGGTGGAAAGACCA 660  
DB 726 CTGGAGAAAGCGCTACCGCTGGGACCGTCAGACAGATTCAAAATGGTGGAAAGACCA 755  
QY 661 TACTTTGTTCAAGCGGTGATACGAGATATATCTACTTCTCTCAGGGAATATGCA 720  
DB 786 TACTTTGTTCAAGCGGTGATACGAGATATATCTACTTCTCTCAGGGAATATGCA 845  
QY 721 GTGAGATATACACCATGGCAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAAAGAT 780  
DB 846 GTGAGATATACACCATGGCAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAAAGAT 905  
QY 781 GATATGGGAGGATCTCAAGAGCTCTGGGAGAACACTGAGCTGCTCTGAGGCGGCG 840  
DB 906 GATATGGGAGGATCTCAAGAGCTCTGGGAGAACACTGAGCTGCTCTGAGGCGGCG 965  
QY 841 TTGAATGCTCATCTCTCCAGACTTCATTTTATTTCANATCTCCAGGACGATTACA 900  
DB 966 TTGAATGCTCATCTCTCCAGACTTCATTTTATTTCANATCTCCAGGACGATTACA 1025  
QY 901 GATGTGATTTCGTATCAACCGCGGTGATGTTGTCCTGGCAACGTTTCTACACTTATAAC 960  
DB 1026 GATGTGATTTCGTATCAACCGCGGTGATGTTGTCCTGGCAACGTTTCTACACTTATAAC 1085  
QY 961 AGCATCCCTGGGTCTGACAGCTGTGCTGATGACATGCTTCACATGCGGAGTGTCTTACT 1020  
DB 1086 AGCATCCCTGGGTCTGACAGCTGTGCTGATGACATGCTTCACATGCGGAGTGTCTTACT 1145  
QY 1021 GGGAGATTTCANGGAACAGAGTCTCTGATTCCAGCTGGACACAGTCTCTGATGAACGA 1080  
DB 1146 GGGAGATTTCANGGAACAGAGTCTCTGATTCCAGCTGGACACAGTCTCTGATGAACGA 1205

QY 1381 GTTCCTAAGCCAGCCAGGCTTGCTGTGCTGGCTCATCTCTCTTAGAACAATATGCAACC 1140  
DB 1206 GTTCCTAAGCCAGCCAGGCTTGCTGTGCTGGCTCATCTCTCTTAGAACAATATGCAACC 1265  
QY 1141 TCCAAATGAGTTCCTCTGATGATACCTGAATCTCATCAAGACGACACCGCTCATGGATGAG 1200  
DB 1266 TCCAAATGAGTTCCTCTGATGATACCTGAATCTCATCAAGACGACACCGCTCATGGATGAG 1325  
QY 1201 GCAGTGGCTTCATCTCTCAACAGGCCATGGTCTCTGAGAACAATGGTCAGATACCGCTT 1260  
DB 1326 GCAGTGGCTTCATCTCTCAACAGGCCATGGTCTCTGAGAACAATGGTCAGATACCGCTT 1385  
QY 1261 ACCAAAATTCAGTGGACACAGCTGCTGGCCATATCAGAAATCACACGTGTTTCTG 1320  
DB 1386 ACCAAAATTCAGTGGACACAGCTGCTGGCCATATCAGAAATCACACGTGTTTCTG 1445  
QY 1321 GGATCAGAGAAGGGAATCATCTTGAAGTTTTTGGCCAGAAATAGGAATAGTGGTTTCTA 1380  
DB 1446 GGATCAGAGAAGGGAATCATCTTGAAGTTTTTGGCCAGAAATAGGAATAGTGGTTTCTA 1505  
QY 1361 AATGACAGGCTTTTCTTGGAGGAGATGAGTGTTCAGAACTCTGAAAATGCACATATGAT 1440  
DB 1506 AATGACAGGCTTTTCTTGGAGGAGATGAGTGTTCAGAACTCTGAAAATGCACATATGAT 1565  
QY 1441 GSAATGCAATACAAAGAAATCATGGCGATCGACAGACAGACAGACAGACAGCTCTCTAT 1500  
DB 1566 GSAATGCAATACAAAGAAATCATGGCGATCGACAGACAGACAGACAGACAGCTCTCTAT 1625  
QY 1501 GTTCGCTCTCTAGCTGTGTGATTAAGGTTCCCTTGGCTGGTGTGAACACATGGAAG 1560  
DB 1626 GTTCGCTCTCTAGCTGTGTGATTAAGGTTCCCTTGGCTGGTGTGAACACATGGAAG 1685  
QY 1561 TGTAAAANAAGCTTATTTCCTCCAGAGACCATATTGCGATGGATTAAGAGAGGTGGT 1620  
DB 1686 TGTAAAANAAGCTTATTTCCTCCAGAGACCATATTGCGATGGATTAAGAGAGGTGGT 1745  
QY 1621 GCCTGCAGCCATTATCACCCAAACAGCAGACAGTACTTTGAGCAGSACATAGAGCGTGC 1680  
DB 1746 GCCTGCAGCCATTATCACCCAAACAGCAGACAGTACTTTGAGCAGSACATAGAGCGTGC 1805  
QY 1681 AATACAGATGGTCTGGGAGCTGTCAANTCTTTGTGGCACTGAATGGSCATTCAGT 1740  
DB 1806 AATACAGATGGTCTGGGAGCTGTCAANTCTTTGTGGCACTGAATGGSCATTCAGT 1853  
QY 1741 TCCCTCTTGCCACACAAACACATCAGATCGACGGCTCAAGAGGCGTATGATCTAGG 1800  
DB 1854 TCCCTCTTGCCACACAAACACATCAGATCGACGGCTCAAGAGGCGTATGATCTAGG 1853  
QY 1801 GGAGGAATGTTGAGTGGAGCTGGAGCATCTGCTGACTCACTTACAGCAGACAGACCTTGGGG 1860  
DB 1854 GGAGGAATGTTGAGTGGAGCTGGAGCATCTGCTGACTCACTTACAGCAGACAGACCTTGGGG 1853  
QY 1861 GCAGTCTTCCCATATATACCAAGACAGAGGAGGATTCGGGAAGTACCTTCAAA 1920  
DB 1854 GCAGTCTTCCCATATATACCAAGACAGAGGAGGATTCGGGAAGTACCTTCAAA 1880  
QY 1921 GGCACAGCAGCTGGTTCCTGCTCAGCTCTTGGCCATTCAGTCTTCCTTGGCTTCGCTC 1980  
DB 1881 GGCACAGCAGCTGGTTCCTGCTCAGCTCTTGGCCATTCAGTCTTCCTTGGCTTCGCTC 1940  
QY 1981 ATGGGGGCGCTCTTCTCGGGCATCACCGTCTACTGCGTCTGTGATCATCGCGCCAAAGAC 2040  
DB 1941 ATGGGGGCGCTCTTCTCGGGCATCACCGTCTACTGCGTCTGTGATCATCGCGCCAAAGAC 2000  
QY 2041 GTGGCTGTGATGACAGSCAGSAGAGSACTCACCCACTGGCCCGGGCTCCATGAGC 2100  
DB 2001 GTGGCTGTGATGACAGSCAGSAGAGSACTCACCCACTGGCCCGGGCTCCATGAGC 2060  
QY 2101 AGGCTCAACAGCTCAAGCGGCTCTTTGGGACACTCAATCAAGACCCAAAGCCGAG 2160  
DB 2061 AGGCTCAACAGCTCAAGCGGCTCTTTGGGACACTCAATCAAGACCCAAAGCCGAG 2120  
QY 2161 GGCATCTCAGCCACTCATGACAAAGGCAAGCTCACCACCTCCCTCCGCAACACGCGCAAG 2220

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 3018)  
 Klostermann.A., Lutz.B., Gertler.F. and Behl.C.  
 The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal tyrosin-like domain  
 J. Biol. Chem. 275 (50), 35647-35653 (2000)  
 235647339  
 10993894  
 2 (bases 1 to 3018)  
 Klostermann.A. and Behl.C.  
 Direct Submission:  
 Submitted (21-JUL-2000) Independent Research Group  
 Neurodegeneration, MPI of Psychiatry, Kraepelinstrasse 2-10, Munich  
 80604, Germany  
 Location/Qualifiers  
 1..3018  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /feature\_type="brain"  
 1..3018  
 /codon\_start=1  
 /product="axon guidance signal SEMA6A1"  
 /protein\_id="AA029494.1"  
 /db\_xref="GI:11093909"  
 /translations="MPSALLLCTLLHCAGAFPEDESEPTISIHGNYTKQYPVFVCH  
 KPGKNTQRTREHLIDQIMINLRNTELYVAARDHYIVYDITDTSHEETIYCKSLIKSRGAQA  
 QVDCTRMKGRKHKECHINIKVLKKNDOTDFVCGTNAFNPSCNTRVUTLEIGDES  
 GMARQDQKHAIVDAFAGKYLSATVDFLAIDAVIYRPGDSPTLRVVKHDSKWLK  
 EYEPVQAVDQDIYFFIEFAIYNTMGVYPPVAVQVCKNDMGGSORVLEKQMTSF  
 LKAVTGRVSGDSGHFVNTLQAVTDVIRINGRDVVLATFSPYNSIPGSAYCAVDMLD  
 LDVFTGRENCEKSPQSTNPVDPDERVPRKPRGCCAGSSLEKVAISNEFPDDTLNI  
 KTHSLMDPAKVIIMNRPFLRMVRYLTKIAVDNAAGPYONHTVVFLENNVYNPKK  
 CSDYDQVRSIMDLDRAGSLYVAFSTCVKLVPGRCHRGKCKTKIASRDYPYCG  
 WVRSGSCAHLSPSLRTEPDITRGNTDGLGDCHNSFVALNGHASSLYPNTTSDSA  
 SRDYESRGGMLDNLNLAEPDPLGAYSSHHQDKKGVIRESKNSQLVPTLT  
 LAIVAILAVYGAAGVGLIYVCDDRDKDAVYVRKKELTTHSRGSSMSYSELGSG  
 FGDKSDKPEALITPLMHGKLATSSNTAKMLKADQHHLDLDTALPTPESTPTLQGG  
 KRPKNRGSREWERQNTINACTKDPMPGSPVITDPLFRASPHISVWYLP.TQGG  
 YQHYVDQPKRSPVQAQMLAQDALEKTKTKEHLSKPNHGVNLVLENDLSLPKPK  
 YQPCASLGPCTISLQSGTGLKLEMOHSSYGLYKRSVPTNSLTRSHOTTILAKRNT  
 NSSNSHLRNQSFGRGPNAPPAPRVDSIQVHSSQPSQAVTVSRQPSLMAVNSLFR  
 SLGATPESLPPVPRKPPAPRVDSIQVHSSQPSQAVTVSRQPSLMAVNSLFR  
 774 a 859 c 744 q 641 t  
 BASE COUNT  
 ORIGIN  
 Query Match 76.1%; Score 2414.6; DB 10; Length 3018;  
 Best Local Similarity 87.1%; Pred. No. 6;  
 Matches 2703; Conservative 0; Mismatches 312; Indels 61; Gaps 2;  
 QY 1 ATGAGGTCAGAGCGCTTGCCTGCTATTATTCACACTGCTACGTTGCTGGCGCTGGTTTC 60  
 DB 2 ATGCGGGCCAGAGCGCTTACTGCTGTCTCCTACACTGCTACCTGCGCGGGCGGGTTTC 60  
 QY 61 CCAGAAGATTCAGAGCAATCAGTATTTCGATCGGCAACTATACAAACAGTATCCGGTG 120  
 DB 61 CCAGAAGATTCGAGCCAAATCAGTATTTCGCA:CGCAACTATACAAACAGTATCCGGTG 120  
 QY 121 TTGTGGGTCACAGGTCAGAGCGGAACACACACAGGACAGGCTGCAGATCCAGATG 180  
 DB 121 TTGTGGGTCACAGGTCAGAGCGGAACACACACAGGACAGGCTGCAGATCCAGATG 180  
 QY 161 AATATGATCATCAAGCAAGCCCTTACATTGCTGCTAGGACCATATTTATACCTGTTCAT 240  
 DB 161 ATCATGATCATCAAGCAAGACCCCTTACGTTGCTGCTCGAGACCATATTTATACCTGTTCAT 240  
 QY 241 ATAGACACATGCACACAGGAAGAAATTTATTGTAGCAAAAACACTGCATGGAATCTAGA 300  
 DB 241 ATAGACACATCCCAACAGAGAAATTTACTGTAGCAAAAACACTGCATGGAATCTAGA 300



301	QY	301	CAGGCCGATGTAGACACATCTCAGCAATGAAGGGAAACATAAAGCATGAGTGGCCACAAC	TTT	360
301	Db	301	CAAGCTGACCTAGACACATCTCAGGATGAAGGGGAAACATAAAGCATGAATGTACAACTTC		360
361	QY	361	ATTAAAGTCTTCTAAAGAAACAGCATGATGCAATGTTTGTCTGTGGAACTAATGCTTTC	420	
361	Db	361	ATTAAGTTCCTCAAGAAAGATGATGATACCTGTTTGTCTGTGGAACTAATGCTTTC	420	
421	QY	421	AACCCCTTCCTGCAGAAACATAAAGAAGGATACATGGAACCAATTCGGGGATGAATTCAGC	480	
421	Db	421	AACCCCTTCCTGCAGAAACATAAGGGTGATACCTGGAACCTTTTGGGAAGAATTCAGC	480	
481	QY	481	GGAATGGCAGATGCCCAATGAATGATCCAAACAAGGCCAACTGTTTCAGATGTTTCAGATCGA	540	
481	Db	481	GGAATGGCAGATGCCCTTAATGATGAGGAAACATGTCNAATCCCTCTGTTGAGATGGA	540	
601	QY	541	AAACTATATCAGCCACACAGTACTGCTCCCTTGCCTATGAGGACATCATTTACCGAGT	600	
601	Db	541	AAACTCTATCCGCTACAGTACTGAGTCTTGTGGCAATGATGGGTGCTATTCACAGAGC	600	
601	QY	601	CTTGGAGAAAGCCCTAACCCAGCCAGCCGCAACACAGCATCAAAAAGGTTGAAGAACCA	660	
601	Db	601	CCCGAGAACAGCCCTACCCACAGCATCTGCAGCATGATCAAGTGAGTTGAAGAGCCG	660	
661	QY	661	TACTTTCTTCAACCCCGTGAATACGGAGATTAATCATCTTCCTTCTCAAGGAAATACCA	720	
661	Db	661	TACTTGTCCAGCCGAGATTAATGGAGCAATATATCACTCTCTTCCAGAGAAATGCA	720	
721	QY	721	GTGGAGTATAACAGCAAGGAAAGTAACTTTCGCAAGAGTCACTCAGTGTCTTAGAAAT	780	
721	Db	721	GTAGAATACCAACTATGGGAAAGTGTCTTCCCTAGGCTTCCTCAGTGTCTTAAGAT	780	
781	QY	781	GATATGGAGGATCTCAAGAGCTCTGAGAAACAGTGGAGCTGCTTCTGGAAGAGGCGG	840	
781	Db	781	GATATGGAGGAGTCTCAGAGATCTCTGGAGAGCAGTGGACATCTCTCTGAAGAGTCCG	840	
841	QY	841	TTGACTGTCTCAGTCTCTCGAGACTCTCATTTTATTTCAACATCTCCAGGCGAGTTACA	900	
841	Db	841	CTGACTGTCTGGTGTCTGAGACTCTCATTTTATTTCAATATACCTCAGGCGAGTTACA	900	
901	QY	901	GATGTGAATCGTATCAAGGGCGCTGATGTGTCTCTGGCAACGTTTCTACACCTTTAA	960	
901	Db	901	GATGTGAATCCGCAATATGCCCCTGATGTGTCTTGGCAACCTTTTCCACACTTATAAC	960	
961	QY	961	AGCATCCCTGGGTCTGCATCTGTGCTATAGCATCTCTGAAATCCCATGTTTATTACT	1020	
961	Db	961	AGCATCCCGAGTCTGCATCTGTGCTATAGCATCTCTGAAATCCCATGTTTATTACT	1020	
1021	QY	1021	GGAGATCTCAAGAAAGAGAGTCTCTGATCTCCACTGGACACACAGTCTCTGATGAACGA	1080	
1021	Db	1021	GGAGGTCTCAAGAAAGAAATCAGTCACTCTACTCTGGACACCCGCTCCAGAGCAACGA	1080	
1081	QY	1081	GTCTCTAAGCCCAAGGCGAGCTGTCTGTGGCTCATCCCTTTAGAAACATATGCAACC	1140	
1081	Db	1081	GTCCCTAAGCCCAAGGCGCAGCTGTCTGTGGCATCTCCCTTTAGAAATATGCAACC	1140	
1141	QY	1141	TCCAAATGAGTCTCCTGATGATACCCCTGAACCTGTATCAAGAGGACCCGCTCATGATGAG	1200	
1141	Db	1141	TCCAAATGAGTTTCCGATGATACCCCTGAACCTGTATCAAGAGGATCCGCTCATGATGAG	1200	
1201	QY	1201	GCAGTGCCTCCATCTTCAACAGGCGATGTTCTCTAGAAACAATGGTCAGATACCCGCTT	1260	
1201	Db	1201	GCAGTGCCTTCCATCTATCAACAGACCTTGTTTCTCTAGAACACAATGTCAGTACCCGCTT	1260	
1261	QY	1261	ACCAAAATTCGATGGGACACAGCTGGTGGGCCATATCAGAAATCAACACTGTGGTTTCTG	1320	
1261	Db	1261	ACCAAAATTCGATGACACACAGCTGGCGGCCCATATCAGAAATCAACACTGTGG	1320	
1321	QY	1321	GGATCAGAGAAAGSNAATCATCTTGAGCTTTTGGCCCAATAGGAATATGCTGTTTCTA	1380	
1313	Db	1313	-----	1312	
1381	QY	1381	AATCAGACGCTTTTCCCTGGAGAGTCACTGTTTACAACCTCTCAAAATACCACTATGAT	1440	

[illegible]









1094 GGGAGGTTCAAGGAGACAGAAATACCCGACATCTTACCTGCAATATACCTGCTTCCAGACGAAGCA 1193  
 QY  
 1091 GTTCTAAGCCAGCCAGGAGTGTGTGCTGGCTGATCCCTCTAGAAAGATATGGAACC 1140  
 DB  
 1154 GTCTCTAAGCCAGGAGGAGGAGTGTGTGCTGGAATATCTCTTGTAGAAATATATGGAACC 1213  
 QY  
 1141 TCCAAATGATCCCTGATATATACCTGATCTATATCAAGAGGACACCGGCTCATGGATGAG 1200  
 QY  
 1214 TCCAAATGATCCCTGATATATACCTGATCTATATCAAGAGGACACCGGCTCATGGATGAG 1273  
 DB  
 1201 GCATGCTCTCATCTTCAACAGGAGATGTTCTGTGAAATATATGTCATATATAGGCTC 1240  
 QY  
 1274 GCATGCTCTCATCTTCAACAGGAGATGTTCTGTGAAATATATGTCATATATAGGCTC 1333  
 DB  
 1261 ACCAAATGTCAGTGGACAGAGTGTGTGGGCAATATCAGAAATATATGTTGTTTCTG 1320  
 QY  
 1334 ACCAAATGTCAGTGGACAGAGTGTGTGGGCAATATCAGAAATATATGTTGTTTCTG 1393  
 DB  
 1321 GCATCAGAGAGGGAATATCTTGAATTTTGGCTGAGATAGAAATATGTTTCTG 1380  
 QY  
 1394 GCATCAGAGAGGGAATATCTTGAATTTTGGCTGAGATAGAAATATGTTTCTG 1453  
 DB  
 1381 AATGACAGCTTTTCTGTGAGAGAGTGTGTGTTTACAGCTTAAATATGAGTATGAT 1440  
 QY  
 1454 AATGACAGCTTTTCTGTGAGAGAGTGTGTGTTTACAGCTTAAATATGAGTATGAT 1513  
 DB  
 1441 GGATTCGAGAGAGGATATGAGGATGAGTGTGAGAGAGAGAGAGAGAGTGTGTGTAT 1500  
 QY  
 1514 GGATTCGAGAGAGGATATGAGGATGAGTGTGAGAGAGAGAGAGAGAGTGTGTGTAT 1573  
 DB  
 1501 GTTGGTTCCTTACCTGTGTATTAAGTGTGCTTGGCTGAGATAGAGAGAGAGAG 1560  
 QY  
 1574 GTTGGTTCCTTACCTGTGTATTAAGTGTGCTTGGCTGAGATAGAGAGAGAGAG 1633  
 DB  
 1561 TGTAAAGAAAGCTGTATGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGT 1620  
 QY  
 1634 TGTAAAGAAAGCTGTATGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGT 1693  
 DB  
 1621 GCTGTCAG 1680  
 QY  
 1694 TCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1753  
 DB  
 1681 AATGACAGCTTTTCTGTGAGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
 QY  
 1754 AATGACAGCTTTTCTGTGAGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1813  
 DB  
 1741 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
 QY  
 1814 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1873  
 DB  
 1801 GAG 1860  
 QY  
 1874 GAG 1933  
 DB  
 1861 GAG 1920  
 QY  
 1934 GAG 1993  
 DB  
 1921 GAG 2000  
 QY  
 1994 AGCAAG 2053  
 DB  
 1981 ATGGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
 QY  
 2054 ATGGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2113  
 DB  
 2041 GTGGCTGTGGTGCAG 2100  
 QY  
 2114 GTGGCTGTGGTGCAG 2173  
 DB  
 2101 AGCTGACCAAG 2160  
 QY  
 2174 AGCTGACCAAG 2233  
 DB

2161 GCATCTCTACGCGCACTCATGCAACGAGCAAGCTGGCCACTCCGGCAACACGCGCAAG 2220  
 QY  
 2234 GCATCTCTACGCGCACTCATGCAACGAGCAAGCTGGCCACTCCGGCAACACGCGCAAG 2293  
 DB  
 2221 ATGCTCATTAAG 2280  
 QY  
 2294 ATGCTCATTAAG 2353  
 DB  
 2281 ACCCAAG 2340  
 QY  
 2354 ACCCAAG 2413  
 DB  
 2341 CAGAACTCATTAAG 2400  
 QY  
 2414 CAGAACTCATTAAG 2473  
 DB  
 2401 AGGAGCT 2460  
 QY  
 2474 AGGAGCT 2533  
 DB  
 2461 AGGAGCT 2517  
 QY  
 2534 AGGAGCT 2593  
 DB  
 2518 CAGATGCGCTCTGAG 2577  
 QY  
 2594 CAGATGCGCTCTGAG 2653  
 DB  
 2578 ACCAG 2608  
 QY  
 2654 AGTAG 2684  
 DB

RESULT 12  
 RD157826  
 LOCUS 2293 bp DNA linear PAT 17-JAN-2003  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION RD157826  
 VERSION BD157826.1 GI:27863584  
 KEYWORDS JP 2002191363-A/12665.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2293)  
 AUTHORS Ota,T., Isoqali,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 TITLE Primer for synthesizing full-length cDNA and use thereof.  
 JOURNAL Patent: JP 2002191363-A 12669 09-JUL-2002;  
 COMMENT HELIX RESEARCH INSTITUTE  
 US Homo sapiens (human)  
 PN JP 2002191363-A/12665  
 PD 09-JUL-2002  
 PE 28-JUL-2009 JP 2002280950  
 PI TOSHIO OTA, TAKAO ISOQALI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORI  
 SAITO,  
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
 PI KEIICHI NAGAI, TETSUJI OTSUKI  
 PC C12N15/09.C07K14/47.C07K16/18.C12N1/15.C12N1/19.C12N1/21.C12N5/ PC  
 10, C12P21/02.C12Q1/68//C12P21/08.G06F17/30.C12N15/00.C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH Key  
 Location/Qualifiers  
 FT CDS location(qualifiers)  
 1..2293  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 579 a 713 c 581 g 420 t  
 BASE COUNT  
 ORIGIN

```

Query Match      59.0%  Score 1824.4;  DB 6;  Length 2293;
Best Local Similarity 99.9%;  Prod. No. 0;
Matches 1825;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

1268	y	TTGCAGTGGACACAGCTGCTGGGGCATATCATAGATACACATGCTGGTATGATGGATCGAG	1327
1	b	TTGGAGTGGACACAGCTGCTGGGGCATATCATAGATACACATGCTGGTATGATGGATCGAG	60
1328	y	AGAAGGGAAATCATCTTCAAGTTTTGGTCAGAAATAGCAAAATAGCTTTCTAAATGACA	1367
61	b	AGAAGGGAAATCATCTTGAAGTTTGGCCACAATAGGAATAGTGGTTTCTAAATGACA	120
1368	y	GCCTTTTCCTGGAGAGATGATGCTTTACACTCTGAATAATGCAGTATGATGGATCGG	1447
121	b	GCCTTTTCCTGGAGAGATGATGCTTACAACTCTGAATAATGCAGTATGATGGATCGG	180
1448	y	AAGACAAAAGGATCATGGCATGCGAGCTGGACACAGCAACACAGCTCTCTGTATGCTGGGT	1507
181	b	AAGACAAAAGGATCATGGCATGCGAGCTGGACAGACAGCAACACCTCTCTGTATGCTGGGT	240
1508	y	TCTCTACCTGTGCTATAAAGTTCGCTTGGCCGTGTGACACATGAGGAATGTA	1567
241	b	TCTCTACCTGTGCTATAAAGTTCGCTTGGCCGTGTGACACATGAGGAATGTA	300
1568	y	AAACCTGTATTCGCTCCAGACAGCCATATCTGTGATGATAAAGCAAGGTGGTGCCTGCA	1627
301	b	AAACCTGTATTCGCTCCAGACAGCCATATCTGTGATGCAAAAGCAAGGTGGTGCCTGCA	360
1628	y	GCCATTTATCACCAACAGACAGCTGACITTTGAGCAGACATAGACGGTGGCAATACAG	1687
361	b	GCCATTTATCACCAACAGACAGCTGACITTTGAGCAGGACATAGACGGTGGCAATACAG	420
1688	y	ATGGTCTGGGAGCTGTCCAAATTCCTTTGGTGCACATGAATGGCATTCGACGTCCCTCT	1747
421	b	ATGGTCTGGGAGCTGTCCAAATTCCTTTGGTGCACATGAATGGCATTCGACGTCCCTCT	480
1748	y	TCCCACTACCAACCATATCAGATTCGACGGGTCTGAAGAGGGGTATGATGTAGGGAGAA	1807
481	b	TCCCACTACCAACCATATCAGATTCGACGGGTCTGAAGAGGGGTATGATGTAGGGAGAA	540
1808	y	TGCTGGACTGGAGCATCTGCTTGACTCAATGTGACACACAGACACCTTTGGGAGCATCT	1867
541	b	TGCTGGACTGGAGCATCTGCTTGACTCAATGTGACACACAGACACCTTTGGGAGCATCT	600
1868	y	CTTCCCATTAATCACCAAGCAAGAGGAGTGTATCGGGAAGTTAGCTCAAGAGCCACG	1927
601	b	CTTCCCATTAATCACCAAGCAAGAGGAGTGTATCGGGAAGTTAGCTCAAGAGCCACG	660
1928	y	ACCAGCTGGTTCGGGTCACTGCTTGGGCAATGGCAGTATGCTTGGCTTCTGTCATGGGG	1987
561	b	ACCAGCTGGTTCGGGTCACTGCTTGGGCAATGGCAGTATGCTTGGCTTCTGTCATGGGG	720
1988	y	CGGCTCTCTCGGGCATCAGCTGCTACTCGCTCTGTATCATCGGGCAAGAGCTGGCTG	2047
721	b	CGGCTCTCTCGGGCATCAGCTGCTACTCGCTCTGTATCATCGGGCAAGAGCTGGCTG	780
2048	y	TGGTGCAGCGCAAGCAAGAGCTTCACCCATCTCGGCGCGGGGGCTCCATCAGCAGCTCA	2107
781	b	TGGTGCAGCGCAAGCAAGAGCTTCACCCATCTCGGCGCGGGGGCTCCATCAGCAGCTCA	840
2108	y	CCAGCTCAGGGCTCTTTTGGGACACTCAATTCANAGACCCCAAGACCCGGAGGCCATCC	2167
841	b	CCAGCTCAGGGCTCTTTTGGGACACTCAATTCANAGACCCCAAGACCCGGAGGCCATCC	900
2168	y	TCACGCCACTCATKCAAAAGGCAAGCTTCGCCATCTCCCGGCAACACGGCAAGATGCTCA	2227
901	b	TCACGCCACTCATKCAAAAGGCAAGCTTCGCCATCTCCCGGCAACACGGCAAGATGCTCA	960
2228	y	TTAAAGACAGACAGCAGCACTGACCTGACGGCTTCGCCACCCCAAGAGTCAACCCCAA	2287
961	b	TTAAAGACAGACAGCAGCACTGACCTGACGGCTTCGCCACCCCAAGAGTCAACCCCAA	1320

QY	2288	CGGTGCAGCAAGAGCGGAGGCCAGCCGCGGACGCCGAGTGGGAGAGGAACACAGAAC	2347
DB	1021		1080
QY	2348	TCATCAATGCTGCACAAAGGACAGCCGCCATGGGCTCCCTGTGATTCGCCAGGACC	2407
DB	1081		1140
QY	2408	TGCCCGTGGCGGCTCCCGCAGGCACATCCCGAGCGTGGTGGTCTCTGCCATCATCGAGC	2467
DB	1141	TGGCCCTTGGCGGCTCCCGCAGGCACATCCCGAGCGTGGTGGTCTCTGCCATCATCGAGC	1200
QY	2468	AGGCTACCAAGCATGAGTACGTGACACAGCCCAAAATAGGCGAGTGGCCAGATGGCG	2527
DB	1201	AGGCTACCAAGCATGAGTACGTGACACAGCCCAAAATAGGCGAGTGGCCAGATGGCG	1260
QY	2528	TGAGGACCAAGCGCGGCACACTGGATATAGACCAATCAAGAGCAATCTTCAGCAGCAAGA	2587
DB	1261	TGAGGACCAAGCGCGGCACACTGGATATAGACCAATCAAGAGCAATCTTCAGCAGCAAGA	1320
QY	2588	GTCCCAACCAAGGGTGAACCTTGTGAGAACCTGGACAGCTGCCGCCAAAGTTCCAC	2647
DB	1321	GTCCCAACCAAGGGTGAACCTTGTGAGAACCTGGACAGCTGCCGCCAAAGTTCCAC	1380
QY	2648	AGGGGAGGCTCTGCTGGTCCCGGGAGCGCTCCCTGTCTCAGACGGTCTTAAGCAAGC	2707
DB	1381	AGGGGAGGCTCTGCTGGTCCCGGGAGCGCTCCCTGTCTCAGACGGTCTTAAGCAAGC	1440
QY	2708	GGCTGGAANTGCACCACTCTCTCTACGGGTTGACTATAGAGGAGGTACCCGACGA	2767
DB	1441	GGCTGGAANTGCACCACTCTCTCTACGGGTTGACTATAGAGGAGGTACCCGACGA	1500
QY	2768	ACTCGCTCAGGAAAGCCACAGGCCACCACTCTCAAAAGAAACAACTAATCTCTCCA	2827
DB	1501	ACTCGCTCAGGAAAGCCACAGGCCACCACTCTCAAAAGAAACAACTAATCTCTCCA	1560
QY	2828	ATTCTCTCAGCTCTCAGAACCAAGCAGCTTGGCAGGGGAGACAAACCGCGCGCGGCC	2887
DB	1561	ATTCTCTCAGCTCTCAGAACCAAGCAGCTTGGCAGGGGAGACAAACCGCGCGCGGCC	1620
QY	2888	CGCAGAGGTGGACCTCCATCAGGTGCACAGCTCCAGGCACTCTGGCAGGCGCTGACTG	2947
DB	1621	CGCAGAGGTGGACCTCCATCAGGTGCACAGCTCCAGGCACTCTGGCAGGCGCTGACTG	1680
QY	2948	CTCGAGGCGAGCTAGGCTCAACGCCATCAACTCAGCTCAGCTGCGGCTGAAGCGTA	3007
DB	1681	CTCGAGGCGAGCTAGGCTCAACGCCATCAACTCAGCTCAGCTGCGGCTGAAGCGTA	1740
QY	3008	CGCCTCTGCTANACCGGAGGATACCCCGCAAGCATCTTGTCTCCCTTTCCACATCCA	3067
DB	1741	CGCCTCTGCTANACCGGAGGATACCCCGCAAGCATCTTGTCTCCCTTTCCACATCCA	1800
QY	3068	TGAAGCCCAATGATGCTGTTACATAA	3093
DB	1801	TGAAGCCCAATGATGCTGTTACATAA	1826
RESULT 13			
AK027471			
LOCUS			
DEFINITION	AK027471	2293 bp	linear
		Homo sapiens cDNA FLJ14565 fis, clone NT2RM4000233, highly similar to Mus musculus semaphorin 7A mRNA.	PRI 01-AUG-2002
ACCESSION	AK027471		
VERSION	AK027471.1	GI:14042170	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1	
		Isogai, T., Ota, T., Hayashi, K., Sugiyama, I., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Watanabe, M., Hosokawa, T., Kaku, Y., Kodaira, H., Kondo, H.,	



Sugawara,M., Takahashi,K., Chiba,Y., Ishida,S., Marukawa,K.,  
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Miyakami,K.,  
Ishii,S., Kawai,Y., Saito,K., Yamamoto,O., J., Wakamatsu,A.,  
Nakamura,Y., Nagahashi,K., Masuko,Y., Nishimura,K. and Iwayama,A., J.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2293)  
Isodai,T. and Otsuka,T.  
Direct Submission  
Submitted (10-MAY-2001) Takao Isodai, Helix Research Institute,  
Genomics Laboratory, 1592-3 Yana, Kisarazu, Chiba 252-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-1975, Fax:81-438-52-3984)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full library construction;  
Research Association for Biotechnology: cDNA library construction;  
5'- & 3'-end one pass sequencing and clone selection; Helix  
Research Institute (supported by Japan Key Technology Center etc.)  
and Department of Virology, Institute of Medical Science,  
University of Tokyo.

## FEATURES

Location/Qualifiers

1..2293  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RM400233"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RM4"  
/note="cloning vector: pUC19FL3-mRNA from uninduced N:2  
neuronal precursor cells."  
138..1826  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB55136.1"  
/db\_xref="GI:14042171"  
/translation="MSVYNSEKCSVDGVEDKRIKGMQLDRASSLSYVAFSTCVTKVPL  
GKERGKKTKCIASRPYCYGKIKEGGACSLSPNSRLTFEIDIERNTDGLDCDN  
SFVALNGHSSLLPTTSDTAQYESRGMIDMKHLLDSPDSTPLGAVSHNQC  
DKGVIRRESYLKHQLVPVTLIAIAVLAVMGAVFSGITVYVCDBRRKDDAVVQR  
KKELTLSRSGMSVTKLFGDTQSKPKPAIITPLMHNGKLATPGNTAKMLIK  
ADQHLDTALPTPESTPLTQOKRPSRSEWERNONLINACTKMPWPMGSPVPTD  
LP-RASPSHPSVVLPIIOGYOHEYDOPKMEVAQMALEDQAILEYKLEHLS  
SKSPRHGVNLVSLPKVPORASLCPGASLSQTLKRLHNMHSSEYGVYDKR  
SYPTLSLHQAHTLKNNTNSHLSRNGFGRDNPFPAPORVDSIQVHSSOP  
SQQAVTVSQPSLNAISLTSGLKRTSLKPDVPPKPSFAPLSISKPNDACT"  
BASE COUNT 579 a 713 c 581 g 420 t

Query Match 59.0% Score 1824.4 DB 3; Length 2293;

Best Local Similarity 99.9% Pred. No. 0;

Matches 1825; Conservative C; Mismatches 1; Indels 0; Gaps 0;

QY 1268 TTCACTGACACAGAGTGTGGGCAATACAGAAATACACAGTGGTGTGGGATGAG 1327

DB 1 TTCACTGACACAGAGTGTGGGCAATACAGAAATACACAGTGGTGTGGGATGAG 60

QY 1328 AGAAGGCAATGATCTGGAAGTTTTCGACAGATAGGAATAGTGGTGTTCATATGCA 1367

DB 61 AGAAGGCAATGATCTGGAAGTTTTCGACAGATAGGAATAGTGGTGTTCATATGCA 120

QY 1386 GCGTTTCTGGAAGACATGATGTTTACAACTCTCAAAATGACGATGATGAGCAJTCG 1447

DB 12 GCGTTTCTGGAAGACATGATGTTTACAACTCTCAAAATGACGATGATGAGCAJTCG 189

QY 1448 AAGACAAAATGATATGGGATGACATGGAATAGACAGATAGATGCTCTGTATGATGCT 1507

DB 181 AAGACAAAATGATATGGGATGACATGGAATAGACAGATAGATGCTCTGTATGATGCT 240

QY 1506 TCTCTACCTGTGTGATAAGTTCCGCTTGGCGCTGTGAAGACATGGAAGTGTAA 1557

DB 241 TCTCTACCTGTGTGATAAGTTCCGCTTGGCGCTGTGAAGACATGGAAGTGTAA 309

QY 1568 AAACCTGTATTCCTCCACAGAGACCCATATTTGTGGATGATTAAGGAAGTGTGGTGTCA 1627

DB 301 AAACCTGTATTCCTCCACAGAGACCCATATTTGTGGATGCATTAAGGAAGTGTGGTGTCA 360

QY 1528 GGCATTTATCAGCCACACACAGACAGCTTTTGTGACAGCATAGAGCTGGCAATACAG 1687

DB 351 GGCATTTATCAGCCACACACAGACAGCTTTTGTGACAGCATAGAGCTGGCAATACAG 420

QY 1598 ATGGTGTGGGAGCTGACAAATTCCTTTGTGGCACTGAATGGGCAATTCGCTCTCT 1747

DB 421 ATGGTGTGGGAGCTGACAAATTCCTTTGTGGCACTGAATGGGCAATTCGCTCTCT 480

QY 1748 TCCCAAGCACACACACATCAGATTCGAGCGGTCAAGAGGATATGAGTCTAGGAGAGAA 1807

DB 481 TCCCAAGCACACACACATCAGATTCGAGCGGTCAAGAGGATATGAGTCTAGGAGAGAA 540

QY 1808 TGTGACATGGAGCATCTGCTTGACCTCACCTGACAGACACAGACCTTTGGGGCASTGT 1867

DB 541 TGTGACATGGAGCATCTGCTTGACCTCACCTGACAGACACAGACCTTTGGGGCASTGT 600

QY 1868 CTTCCTATATCAGACACAGAGAGGAGTGTATTCGGAAAGTTACCTCAAGAGCCACG 1927

DB 601 CTTCCTATATCAGACACAGAGAGGAGTGTATTCGGAAAGTTACCTCAAGAGCCACG 660

QY 1928 ACCAGCTGTTCCTGACACCTCTTGGCATTCAGTCTCATCTGCTTTCTGTCATGGGG 1987

DB 661 ACCAGCTGTTCCTGACACCTCTTGGCATTCAGTCTCATCTGCTTTCTGTCATGGGG 720

QY 1988 CCGTCTTCGGGACATCACCGTCTACTGCTGTGTAATCATCGGCGCAAGACGTGGCTG 2047

DB 721 CCGTCTTCGGGACATCACCGTCTACTGCTGTGTAATCATCGGCGCAAGACGTGGCTG 780

QY 2048 TGTGACAGCGCAAGGAGAGAGGAGTCAACCCATCGCGCGGGGCTCCATAGCAGCTCA 2107

DB 781 TGTGACAGCGCAAGGAGAGAGGAGTCAACCCATCGCGCGGGGCTCCATAGCAGCTCA 840

QY 2108 CCAAGCTCAGCGGCTCTTTTGGGACACTCAATCAAAAGACCCAAAGCGGAGGCCATCC 2167

DB 841 CCAAGCTCAGCGGCTCTTTTGGGACACTCAATCAAAAGACCCAAAGCGGAGGCCATCC 900

QY 2168 TCACGGACCTCATGACACAGCGCAAGCTGCGCAGTCCCGGCAACACGCGCAAGATGTCA 2227

DB 901 TCACGGACCTCATGACACAGCGCAAGCTGCGCAGTCCCGGCAACACGCGCAAGATGTCA 960

QY 2228 TTAAGCAGCAGCAGCAGCAGCTGAGCTTACCGGCTTACCGGCTTCCACACAGAGTCAACCCCAA 2287

DB 961 TTAAGCAGCAGCAGCAGCAGCTGAGCTTACCGGCTTACCGGCTTCCACACAGAGTCAACCCCAA 1020

QY 2288 CGCTGACGAGAGCGAAGCCCGCGGAGCGCGGAGTGGGAGAGAGCAGAGAC 2347

DB 1021 CGCTGACGAGAGCGAAGCCCGCGGAGCGCGGAGTGGGAGAGAGCAGAGAC 1080

QY 2348 TCATCAATGCTGTACAAAGACATGCGCCCAATGGGCTTCCCTGTGATTCGAGGAGC 2407

DB 1081 TCATCAATGCTGTACAAAGACATGCGCCCAATGGGCTTCCCTGTGATTCGAGGAGC 1140

QY 2408 TSCGCTCGGCGGCTTCCCGACGACATCCCGAGTGTGTGGTGTGCTGCCATTCAGCAGC 2467

DB 1141 TSCGCTCGGCGGCTTCCCGACGACATCCCGAGTGTGTGGTGTGCTGCCATTCAGCAGC 1200

QY 2468 AGGCTTACGAGCATGATGATGAGTGGACACCCAAATGAGCGAGTGGCGGAGTGGCGG 2527

DB 1201 AGGCTTACGAGCATGATGATGAGTGGACACCCAAATGAGCGAGTGGCGGAGTGGCGG 1260

QY 2528 TGAAGACAGCGCGGACACTGGAGTATAGACCATCAAGAACATCTCAGCAGCAGAA 2587

DB 1261 TGAAGACAGCGCGGACACTGGAGTATAGACCATCAAGAACATCTCAGCAGCAGAA 1320

QY 2588 GTCCCAACCAATGGGTAACCTTGTGGAACTGACAGACCTGACAGCTGCCCGCAAGTTCAC 2647

DB 1321 GTCCCAACCAATGGGTAACCTTGTGGAACTGACAGACCTGACAGCTGCCCGCAAGTTCAC 1380

QY 2648 AGCGGAGAGCTTCCCTGGGTCGCGCGGAGCGCTCCCTGTCTCAGACCGGTCTTAAGCAAC 2707





QY 1770 TTGAGGCTCAGAGGGGTAAGTCTAGGAGGGAATGCTGGACTGGAGCACTGCT 1829  
DB 421 TTGAGGCTCAGAGGGGTAAGTCTAGGAGGGAATGCTGGACTGGAGCACTGCT 480  
QY 1830 TGACTCACTGACACAGAGAGCTTGGGGCACTGTCTTCCCAATAATCACCAAGACAA 1889  
DB 481 TGACTCACTGACACAGAGAGCTTGGGGCACTGTCTTCCCAATAATCACCAAGACAA 540  
QY 1890 GAAGGAGTGTNTCGGAAAGTTACTCTAAAGCCACAGACAGCTGGTTCCGCTCACCT 1949  
DB 541 GAAGGAGTGTNTCGGAAAGTTACTCTAAAGCCACAGACAGCTGGTTCCGCTCACCT 600  
QY 1950 CTGGGATTTGAGTCATCTGGCTTCTGCTATGAGTGGCTCTCTGGGATCACCT 2009  
DB 601 CTGGGATTTGAGTCATCTGGCTTCTGCTATGAGTGGCTCTCTGGGATCACCT 660  
QY 2010 CTACCTGGCTGTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAGAGAA 2069  
DB 661 CTACCTGGCTGTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAGAGAA 720  
QY 2070 GTTACCCACTGCGGCGGGGCTCATGAGCAGCTGTGATGAGTGTGCTGTGCTGTG 2129  
DB 721 GTTACCCACTGCGGCGGGGCTCATGAGCAGCTGTGATGAGTGTGCTGTGCTGTG 780  
QY 2130 GGACACTCAATCAAGAGAGCCAAAGCCGAGGCGCTGCTGACGCGCTCATGTACAA 2189  
DB 781 GGACACTCAATCAAGAGAGCCAAAGCCGAGGCGCTGCTGACGCGCTCATGTACAA 840  
QY 2190 CAAGCTCGGCTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAGAGAA 2249  
DB 841 CAAGCTCGGCTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAGAGAA 900  
QY 2250 GGACCTGAGTGGGCTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAGAG 2309  
DB 961 GGACCTGAGTGGGCTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAGAG 960  
QY 2310 CAGCGGCGGAGCGGAGCGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2369  
DB 961 CAGCGGCGGAGCGGAGCGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 2370 CATGGCGGCTGAGGCTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAG 2429  
DB 1021 CATGGCGGCTGAGGCTGTGATCATCGGCGGAAGAGCTGTCTGTGCAAGCCAAAG 1080  
QY 2430 CCACATCGGAG 2489  
DB 1081 CCACATCGGAG 1140  
QY 2490 GGACAG 2549  
DB 1141 GGACAG 1200  
QY 2550 GGAGTATAG 2609  
DB 1201 GGAGTATAG 1260  
QY 2610 TGTGGAG 2669  
DB 1261 TGTGGAG 1320  
QY 2670 CCGGAG 2729  
DB 1321 CCGGAG 1380  
QY 2730 TTCTAGGAG 2789  
DB 1381 TTCTAGGAG 1440  
QY 2790 GGCGAG 2849  
DB 1441 GGCGAG 1500

QY 2850 CCAGAGCTTTGGCAGGAGAGACAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2909  
DB 1501 CCAGAGCTTTGGCAGGAGAGACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560  
QY 2910 GGTGCACAGCTTCCAGAGCATCTTGGCCAGGCGGTGACTGTCTCGAGGAGCGCCAG 2969  
DB 1561 GGTGCACAGCTTCCAGAGCATCTTGGCCAGGCGGTGACTGTCTCGAGGAGCGCCAG 1620  
QY 2970 GCGCTACAGCTCAGTACAGAGGTGCGGCTGAGAGGTACGCGCTCGCTAAAGCGGAG 3029  
DB 1621 GCGCTACAGCTCAGTACAGAGGTGCGGCTGAGAGGTACGCGCTCGCTAAAGCGGAG 1680  
QY 3030 ACCCGCCAAAGCATCTCTTGTCTCCCTTTCCATCTCATGAGCGCCAAATGATSGT 3089  
DB 1681 ACCCGCCAAAGCATCTCTTGTCTCCCTTTCCATCTCATGAGCGCCAAATGATSGT 1740  
QY 3090 ATAA 3093  
DB 1741 ATAA 1744

Search completed: September 30, 2003, 20:16:09  
Job time : 11083.7 secs





```

485 -----EPTAISAVELSTYCCQYIGSTAGVCLPHRCQIVYKACAECCCLAPRYCAM--- 537
538 EGACSHLSPNS-RLTFECCIERCQNTDGGCD-----HNSFVALNGHS--SSLPRSTTSD 590
539 DGSACSRYPPTAKERTRCQIRNG--DPCTHCSDLHNDH---HGHSPERIIVGVENS 592
591 STAQCYEYERGERGMLKHLKLDLSPSTDPGAVSSNNHQKGVIRRES 637
593 TEJECSPKORALVYVQFORREERKEE--RVDDHIIITCQGLLRS 638

RESULT 4
49423
Semaphorin I precursor - beetle (Tribolium confusum)
Species: Tribolium confusum
Date: 19-Mar-1997 #sequence_revisor: 19-Mar-1997 #text_change 07-May-1999
Accession: A49423
Colodkin, A.L., Matthes, D.J., Goodman, C.S.
ell 75, 1389-1399, 1993
Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
Reference number: A49423; MUID:94294332; PMID:6269517
Accession: A49423
Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
Molecule type: mRNA
Residues: 1-711 <KOL>
Cross-references: GB125390

Query Match 15.5%; Score 845.5; DB 2; Length 711;
Best Local Similarity 34.3%; Prod. No. 1-46-48;
Matches 230; Conservative 110; Mismatches 270; Indels 165; Gaps 30;

12 LLEFACAGPEDEGEPI-----SISHNNYKQVVFVCHKPGDNTTORHRLDQIMNICH- 65
12 LIALCHAMPOSSSKLINHKSVESKFT-----GNATFCH-----FVING 54
66 --GTLIAARDHLYTVDI--CTSHIEIYCKLTKWSPQADVCTGVKXKKKDECHNIK 122
55 DETSILVCGNRVYNSIFDLERK-----GGIDFSSDAHSCILCKKTDCQNVIR 110
123 VLLKNDLALFVCGTNAPNSORNYMOTLEFFGD--FFSGVACQPYDAKHNALFADGK 191
111 ILYSSRGMVIGCTNSYPLCPTVAFKGGKYLVEKEVEGIGLCTVNPBNSTSVYNGQ 170
182 LYSATVTCFLADAVYRSLGSPFLATVKKHKKLKHYPYCAVQYQYVFFREELAY 241
171 LFSATVAFSGGGLAYR-----PORTLSDLQNLNAPFVNSVAYGVYIEFFYPFTAV 225
242 EYNTMGKVFVFRVAQVCKDMGSSQVLEKQNTSFLKARLNSVPCDSHFYFNILCAVTC 301
226 EYVNCQKVYSRVARVCKDKGFGHSSRDR--NTSFLKARLNSISIEYPPYDFCTSTSD 284
302 VI---RINGRD--VYLAFTFTPNYSIPGSVAVCAVDYLDIASVTFGRFEKQSPDS--WTPVP 357
285 IVEGRVNSDDSKIIYGLITPVAIGSALCAVQVADILRVFESEFKQETIKSNKLPVP 344
358 DERVPVPRCCAGSSSLRYATSNFFPDTLNFKTHP--YDSEAPSIENRPFALTNVR 417
345 QNLVPEPRPQQVCRSRI-----LFDKNVNFILKTHSLMED--VPALFGKPLVYRVSUQ 395
418 YRLTK-ADVTAAAGFYQH--TVVFLSEKGIILKKFLAR--GNSGFLNDSLFLEMSVNSE 475
396 YRFATVDFQVKTINQYLDVLYIGTDGKVLK-----IKVP----- 429
476 KSYDGVEDKRINGXQLDRASSLSLYVAFTCV-----IKVP----- 511
430 -----AVNPKRHAALLYKRYTSVHPHGAPVKQLKAPGKVVVWGKDEIR 478
512 ---LGRCHRGKCKKCIASRDYPCWKIKEGGACSHLSPNSRLTP-ECQIERGN----- 561
479 LAKNHNCASTRC-KDVELQDPHICADAKQNLCVS--DTVTSYRFL--CDVWRGCDNKQWS 537
562 --TDGLGDCCHNSFVALNGHSSLSLPTTSDSTAQNGYBSRGMLQWKILLDSPSTCPD 519

```

```

539 PCTCKKTVTKNK-----PSEVENET-----NSIDKDL-----DSSCPL 572
620 GAVSSNNHQKGVIRRESYKGG---HDQLVPVT--LLAIATLAFVMA---VFSGITVY 671
573 IKTGLDSDSDCDPV--SENSIGGCAVQQLVIVTAGTLHIVVWVSVIGLPSMYLSLSVP 631
672 CV---CHRRKQVAVVORKE--KELTHSRRGNS--SVTKLSGLFGTQSKDKPKPBA--LTL 726
632 AKFHSDSQVPEAPF--ECHNHLELSANQTYCTPRANKAVNLVVKVSSSTPRPKNDLV 691
727 XNKGKLTATPONTAKY 741
692 SKOLNIASDGLQXI 766

RESULT 5
148747
Semaphorin D - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revisor: 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: 148747
R:Buschell, A.W., Adams, R.H., Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: 148744; MUID:95267431; PMID:7748561
A:Accession: 148747
A:Status: preliminary; translated from GB/EXBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-772 <RES>
A:Cross-references: EXBL:X85993; MUID:9854329; PIDN:CAAS1985.1; PID:9854330
C:Genetic:
A:Gene: semD
C:Superfamily: semaphorin

Query Match 15.5%; Score 942.5; DB 2; Length 772;
Best Local Similarity 34.3%; Prod. No. 2-6e-48;
Matches 202; Conservative 85; Mismatches 225; Indels 73; Gaps 16;

QY 47 GRNTQHRKLCQIMVINGT-----LYIAARDHIYVDIET 83
DB 25 GKNVPRFLKSYEMLESNNVITFNGLSKSSVHTFLDSEERSPLYVGAKCHIFSN-- 82
QY 84 SHTSEYCSKLTWKSRCQADVCTCKMKKH--XDCNFIKVLKKYDDALFVCGTNAEP 142
DB 83 ---VNKDPKIVPVSYRDECKWAGKDLKCANFIKYLEANYNOTHLVACGTAFHP 139
QY 143 SCR-----NYKMDTLEFFGDEF--SGVACQPYDAKHNALFADGKLYSATVTCFLADA 195
DB 140 ICTVIEVGHHPEDNI FKLQDSHFENKRGKSPYDPKLLTASLLIDGELYSGTAADFVGRDF 199
QY 106 VYRSLGSESTLTGVKDSKWLKEPYVQA-----VYGDVLYIEFFREIAVEYTMGR 246
DB 200 AIFRTLDGHHPIRTGCHDSKMLNDPFIKAHLIPESDNPEDCKVYFFRENA--GGSHSK 259
QY 249 VYFVPAQVCKDMGSSQVLEKQNTSFLKARLNSVPCG-----DSHFYNILCAVTVIR 304
DB 260 ATRIRGQICNDPFGG--HRSLVNKNITFLKARLICVPGNGIDTHF-----DELQVPL 313
QY 305 INGRD---VYLAFTFTPNYSIPGSVAVCAVDYLDIASVTFGRFEKQSPDSTWTPVDER 360
DB 314 MNSKCPKNPIYGVFTTSSNIFKGSVAVCYMSDVRVFLGPPVAHRDGGNYQWVPV--QGR 372
QY 361 VEPKPRCCAGSSSLRYATSNFFPDTLNFKTHPLMDEAVPSIENRPFALTNVYRL 420
DB 373 VYVPRPTCP--SKTFGCFDSTKGLPDDVITFGSRHPAYNPVPF--INNRP--MIKTQVNYQF 431
QY 421 TKAVDPTAAGPYQNETVFLGSEKGIILKFLARIGNSGFLNDSLFLEMSVNSKCSVD 480
DB 432 TQIVDVRVDEDCQYQWYFGLT--DVTGVLVKVVVPKET--HDLLEVLSEMTVFR----- 484
QY 481 GVEDKR--MGQLDRASSLSLYVAFTCVIKVPLGRCHRGKCKKTC--ASRDPYCGK--KEGG 540

```



b 485 --BPTTISAVELSTKQQLIGSTAGVAGQPLHKCDIYKACAECCIAEIPYCAW--DGS 540  
y 541 ACSHLSPNS-PLTFEPIERGTGGLDCHNSFVALNGHSSSL 542  
b 543 SCSEYPTTASATACQDING--DPLTHCSLDLHCHNHGPFSL 541  
RESULT 6  
49369  
ciliapsin - chicken  
;Species: Gallus gallus (chicken)  
;Date: 07-Apr-1994 #sequence\_revision: 07-Apr-1994 #text\_change: 24-Sep-1994  
;Accession: A49069  
;Uco: F.; Raible, D.; Raper, C.A.  
ell 75, 217-227, 1993  
;Title: Ciliapsin: a protein in brain that induces the collapse and paralysis of neurons  
;Reference number: A49069; MUID:8402978; PID:8402978  
;Accession: A49069  
;Status: preliminary; not compared with conceptual translation  
;Molecule type: mRNA  
;Residues: 1-772 <Uco>  
;Cross-references: GB:06228; MID:34.0078; PID:AA090938.1; PID:8402978  
;Superfamily: semaphorin  
Query Match 15.4%; Score 839.5; DB 2; Length 772;  
Best local similarity 35.2%; Pred. No. 41e-48;  
Matches 194; Conservative 97; Mismatches 220; Indels 57; Gaps 17;  
/ 68 LYAARDHIYVDIDTSHTREIYCSKALTWKSPQADVOTCRMKGKDECHNFVLLK 126  
o 69 LYVAKDHIESFL-----VNIKEYQKIWVSVSHSRDECKWAGKILRECANFIKVLK 123  
y 127 KNDALFVCGTNAPNSCR-----NYKMTLEPPGDEF--SCVAPGYDAKANVALPAD 179  
b 124 YNQHETACGTGAFHPCITVIEVSHPECFIFMEDSHFENGKGYDPKLLTAGLLVD 183  
y 180 GKLYSATVTFPLADAVIYRSLGSEPLRTVKDQSKWKEPFYVQA-----VDYGVY 232  
b 184 GELYSGLAAQFYGRDPAFRTLGHHPIRTEQDSRWLADPRISALIFESDNPEDDK 243  
y 233 YFPRETAVENTMGKVVFPVAVQVCKVGGGQGVLEKQWTSFLKARLNCSPG----C 288  
b 244 YFFRENAIDGHTGKATHARIGQCKNDEGG-HRSVYKWTTFELKARLCSVEGPGWGD 302  
y 289 SHEYENILQAVTVIRNGSD---VVLATFSTPYNSIPGSAYCAVDMLDIASVFTGRK 344  
b 303 THF-----CELCQVFLNSKQPKPIYGVFTSSNIFKGSAYCAVSKTETVRRVLPSPA 357  
y 345 EQKSPSTWTFVDERYKPRPGCCAGSSSLERYATNSNEFPDITLNFIKTHPLVQDAVPS 404  
b 358 HRDGPYQWVPY-QGRVPYPRPGTGP-SKTFGGFDSTKCLPDEVITFANRHPAMWKEVFP 415  
y 405 IFNEPWLRTMVRVRLTKAVDTAAGPYQHTVVFVSEKGIILKEARLGNKSGFLNDSJ 464  
b 416 INSRPIKIDVYCFCTQVVDVRVADGGQIDWPFGTGIDGLVLYKVSIPKETHIEIEV 475  
y 465 FLEMMSVYNSKESYGVDEKRIKNGMLDPAASSLVVAFSTQVYKVLQRCEREGKCKXT 524  
b 476 LLEBYTFV-----EPTVLSAKMISTKQQLVGSATGVSQFLKCDYVGVKAAE 526  
y 525 CIARDPYGMKKGKACSLSPNS-RLTTEQIERNTGGLDCHNSFVALNGHSSSL 583  
b 527 CCLARDPYCAW--DGSSCSRYFPTAKARTRQDIENG--DPLTHCSL-----LQHCHN... 575  
y 584 PSTTTSQSTAGQGVESRGMLD 605  
b 576 PSQVLEKXIYGVRSSTFUS 597

RESULT 7  
5928  
semaphorin F precursor - human  
;Species: Homo sapiens (man)  
;Date: 21-Dec-1996 #sequence\_revision: 06-Jun-1997 #text\_change: 24-Sep-1999  
;Accession: G01856  
;Sex: M, Y.

C:Date: 10-Apr-1998 #sequence\_revision: 08-May-1998 #text\_change: 17-Nov-2000  
C:Accession: J05928  
R:Stammers, A.D.; Plesch, A.W.; McPherson, C.D.; Overhauser, J.; Lovett, M.  
Biophys. Res. Commun. 242, 665-691, 1998  
A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat band.  
A:Reference number: CCS928; MUID:9825554; PMID:9464278  
A:Accession: J05928  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1074 <SIM>  
A:Cross-references: GB:052840; MID:9272583; PID:AA009473.1; PID:9272584  
A:Experimental source: brain  
A:Comment: This protein disrupts normal brain development and leads to some of the feat  
C:Genetics:  
A:Gene: semaF  
C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology  
F1-10/Domain: signal sequence #status predicted <SIG>  
F30-533/Domain: semaphorin #status predicted <SE>  
F340-396/Domain: thrombospondin type 1 repeat homology <THR3>  
F371-993/Domain: transmembrane #status predicted <TM>  
Query Match 15.2%; Score 930; DB 2; Length 1074;  
Best local similarity 36.2%; Pred. No. 3e-47;  
Matches 199; Conservative 90; Mismatches 211; Indels 50; Gaps 17;  
QY 68 LYAARDHIYVDIDTSHTREIYCSKALTWKSPQADVOTCRMKGKDECHNFVLLK 127  
DB 70 LYVAKDHIESFLQ-----ECLSLQAVWECCBATKACYSKSKSCQNYRVLVAV 123  
QY 128 NDDALFVCGTNAPNSCRNYKMTLEPPGDEFGMARCPYDAKFAVALF-ADGKIYSAT 186  
DB 124 GGDLFTCGTNAPTFVCTNRSLSNLAIEHDIQISGMARCPYSPHNSFALLTAGELVAA 193  
QY 187 VTDLAIDAVIYRSLGSEPLRTVKDQSKWKEPFYVQAVDYGDIYFFPRE-AVENTW 246  
DB 184 AMDFGDEPAIYRSLGLPLRLTAQYNSKWLNEPNVSSYDIQNTYFFPRENAVEED-C 242  
QY 247 GKVVFPVAVQVCKNDGSSQGVLEKQWTSFLKARLNCSPVGDSHFVFNILQAVTVIRIN 306  
DB 243 GKTVPRAARVCKNDIGG-APLEDTWTFNKLRLNCSRGVEVFFYINELQSTFFLPEL- 300  
QY 307 GRVVLATFSTPYNSIPGSAYCAVDMLDIASVFTGRFKEQKSEDSWTVPEDERPKPRP 366  
DB 301 -DLIYCFITVNSIAASAVCVFNLSAQAATGSGPKYQVNSRSAMLPYEN--PNPHF 355  
QY 367 CCCAGSSSLERYATNSNEFPDITLNFIKTHPLMDEAVPSIPNRFPLCTMVPYELTAVD 426  
DB 356 QCGTVDDGLYVNLTERNL-QDAOKFLVH-----EVVQPVTVVPSFMEDNSRF--SHVAVD 438  
QY 427 TAAGPYQWHTVVFVSEKGIILKPLARLGNKSGFLNDSLP-EEMSVYNSKESYGVDECKR 486  
DB 409 VVQGREALVHIIVLACDVGTKKVRVPLNQT---SSCCLEELIELEPERR-----REP 458  
QY 487 IMGVQLDRASSLVVAFSTQVYKVLQRCERHCKKCTCIASRDPYCGWKEGACSHLS 546  
DB 489 IRSQILHSQSVLFVGLPEHWKIPKRCQFY-RTSTCIGAGQDPYCGMDVWKKCTSL 57  
QY 547 PMSRLT-PEGDIERGNTDGLDCHNSFVALNGHSSSLPST--TSDSTAGQGVESRGGM 603  
DB 518 ESLSMTQWESLSA-----CPTNLTVDGHPGVMSPTPCTHTDGSV-----GSC 563  
QY 604 LKMKHLLDSP 613  
DB 564 LCRTESSDSP 573

RESULT 8  
501856  
semaphorin F - human  
;Species: Homo sapiens (man)  
;Date: 21-Dec-1996 #sequence\_revision: 06-Jun-1997 #text\_change: 24-Sep-1999  
;Accession: G01856  
;Sex: M, Y.



```

127 PESONPEDDKVYFFRENAIDGHSRGAATHAIGQ::CKNDFFGG-HSLVKNKNTTFLKARL 185
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

228 PAYKERVYFNFREIASEAIDNNEEPQIYARVARVCKNDKGA-RPANERTVSYLKARLNC 286
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

RESULT 12

C49423

seraphorin 11 precursor - fruit fly (*Drosophila melanogaster*)C:Species: *Drosophila melanogaster*

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 07-May-1999

C:Accession: C49423

R:Kolidaki, A.L.; Matthews, D.J.; Goodman, C.S.

Cell 75, 1389-1399, 1993

A:Title: The Seraphorin genes encode a family of transmembrane and secreted growth cone

A:Reference number: A49423; MUID:94094332; PMID:8269517

A:Accession: C49423

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A:Molecule type: mRNA

A:Residues: 1-724 &lt;KOL&gt;

A:Cross-references: GB:LAF083

C:Genetics:

A:Gene: sera 11

A:Cross-references: FlyBase:FBgn001260

C:Superfamily: seraphorin

Query Match: 14.3%; Score 779.5; DB 2; Length 724;

Best Local Similarity: 33.5%; Pred. No. 3.9e-44;

Matches 193; Conservative 99; Mismatches 199; Indels 85; Gaps 22;

CQ 33 NYCKQYVVFVGHKKPQNTTCQRHLD-----ICMININ-----GTLYAARDHYTV 79

CQ 32 NFVYERPCTGNDQGNKNGVGHGADHVEFNGKLYVYRTFFNEDRDTLYVGANDRIVRV 91

CQ 80 DIDCSHTBEIYCSK-KLWNKSPQADVDTCRMKGKHK-DECHNF-KVLLKND-DALFVCG 136

CQ 92 NQNTISSN--CNROAINLEPTRODVSCVSKSQIPDCQNHVIVCSMDQGRLLVCG 149

CQ 137 TNAFNFSGRN-----KNDTLEFFGDFSGVACPCYDAKHAHVAFADG----- 180

```
150 TWANRP--KCVYIVANLTHLPSEVIGV---LGIACPYDPLDNLSTNIYVENONPGGL 204
151 KLYSATVDFLAIDAVYRS-----LGSPTRTVKXDSKWLKTPYFVQAVDYGQY 231
205 PGLYSGTNAEFTKATVAFRTDYNTSARLEKYEKFKTLKYLKVSCKLDRPNFVGSFDIGY 264
232 IYFFREJAVEVNTMGVYVPRVAQVCKXVDSGSORVLEKQKTSFKARLDCSPGDSHP 291
265 VIFFREJAVEVINGKAVYSIAIRVCKXDVGG-KMLAHKWATYCKARLDCSISGEPP 323
292 YNMLCAVTDVIRINGROVVLATFTPPNYSIGSAVCAYDMLCIAVSTIGFKECKSPDS 351
324 YFNEQSVYCLPSDKSR--FFATFTTSTNGLGSAVCSFHINEICAAENGKFKQSSNS 381
352 TWTPVDERVPRPGCCAGSSSLERYATSNFFPDTLNFIKTHLPXDEAVPSIPNRF-W 410
382 AMLPVJNSVPRPGTCVNDTS-----NLPDVLNFIKTHLPXDKAVNHEHNPVY 433
411 PLRTVVRVRLT-KIAVDTAAGPYQKHTVFLGSEKGIILKFLARIGNSGRNDLSFLSE 469
434 YKRDVFTKLVDKRIIDL---NRYIVYVGTNLSRIYKLVQYRNGESLSKLDLIFE 490
469 MSVYSEKCSYGVEDKRIKNGVQLDRASSLSVAFSTCVIKVFLGRC-ERRSKXKTCIA 527
491 VA-----PNEALVMEISQTRKSLVIGTCHRIKQDLANCKRRVNDQPR-CV- 536
528 SDPPCGMIKEGAGSHSPNRLTFEODIERGNTD 563
537 RRPYCGMDKEANTCRPY-----ELLCQVANEHSD 567

RESULT 13
Accession: G02173
Species: Homo sapiens (man)
Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
Accession: G02173
Submitted to the EMBL Data Library, October 1995
Reference number: G02175
Status: preliminary; translated from GB/EMBL/DDBJ
Molecule type: mRNA
Residues: 1-751 <NAV>
Cross-references: EMBL:U02276; NID:G-061150; PIRN:AAB:6276.1; PDB:0C1351
Superfamily: semaphorin

Query Match 14.1%; Score 769; DB 2; Length 751;
Best Local Similarity 31.6%; Pred. No. 2.4e-43;
Matches 191; Conservative 109; Y-smatches 207; Indels 98; Gaps 22;

6 LLYFTLLHFAGAGPESEPISSHGNTKCYPVFVGHKPGRNTTQRRHLDIQVIMI----- 61
6 LLYWASLTLGAWFSEFTQD-----ELP---ATPRVR-SFKELKATG 43
62 -----YIMNGT-----LYIAARSHVTVDID-SHTEEIYCSKKLTKWSQA 102
44 TAFHFNFLNTTCDYILLKDESDHDMYVGSQDYVLSLDHINREPLI-----IFWASQP 99
103 DVDTCRMKAGX-KDCHNFIKVLKKNDDALFVCGTNAPNFSC-----RNYKMDTC 152
100 RIECVLSGKVNCGEGFVRLICPKNRTHLVCGTGAYNPCTYVNRGRACDY-IFYL 158
153 EFGDFGSMARCPYDAKHANVA-FADGKJYSATVTFELACAVIYRSLGESPTLRTYKH 212
159 EPERLE-SGKGCPCYDPKLDTASA-NBELVAGVVDPMGTDAALFRTLGKATRTDQY 217
213 DSKWLKPEYFVQA-----VQYGDYIYFFREIAVYVNTNKKVFRVAQVCKXDGSSQ 266
218 NSRMLNPSFFHAELIPSAENDKLLYFFERSAE-AQSPAVYARIGRICLNDGGHC 276
267 RVLEKQWTSFLKARLNCSPVGDG--HFVFNILQAV-----TDVIRINGROVVLATFTSTPN 320
```

```
277 CLONK-WSTFKARLVCSVPGDGI-EHFDLQGVFQCTQDVR---NPIVIVVFTSSGS 332
321 SFGSAVCAYDMLCIAVFTGRFKECKSPDSTWTPVDERVPKPRGCCAGSSSLERYAT 380
333 VFRGSAVCVSVADISRMVFNSTFAHKEGPNYQMPF-SGKMPPRPGTCGGFTFPMKS 391
381 SNFPDQTLNFIKTHLPXDEAVPSIPNRPKFLTNVYRLTKIA-----VDTAAGPYQNHV 437
392 TKDYPDQVIMFMSHPLXACAVYFLQRRPLVAVITGAPYRLTTIAVGVCSADGRYE--V 448
438 VFLGSEKGIILKFLARIGNSGFJNCSLFEEMSVYNSEKCSYGVEDKRIKNGVQLDRASS 497
449 LFLGTDRGTQVKVIV-LPKDQOEVEELMLEVEVFK-----DPAPVKTWISSKQ 498
496 SJVAFSTCVIKVFLGRCERRHGCKKTCIASRDPYCGWIKESGACSHLSPNS-RLTFEQD 536
499 QLYVASAVGVTHLSHRCCAYGAACADCCJARDPYCAW--DQACSRYTASSKRRSRQD 556
557 IERGN 561
557 VRHGN 561

RESULT 14
semaphorin E - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1995 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: U48748
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: U48748; MUID:95267431; FYID:7748561
A:Accession: U48748
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-751 <RES>
A:Cross-references: EMBL:X85994; NID:G854333; PIRN:CAA59986.1; PDB:G854332
C:Genetics:
A:Gene: sem-E
C:Superfamily: semaphorin

Query Match 13.5%; Score 737; DB 2; Length 751;
Best Local Similarity 31.3%; Pred. No. 2.9e-42;
Matches 187; Conservative 103; Y-smatches 234; Indels 74; Gaps 18;

24 SEPTSSHGNTKCYPVFVGHKPGRNTTQRRHLDIQVIMI-----XNGTLVIAARSHVTVDI 91
42 SEVFLSH-----CQLDYRILLMDSEDCDRIYVGSKHLSLN- 79
62 DTSHTREIYCSKKLTKWSQADYDTCRMKAGXKDE-CHNFIKVLKKNDDALFVCGTNAP 140
90 KNISQELP-----SVFNPASTKYVECKMAGKQPTHGCGNFRVVIQTFNRTHLVYCGSAP 135
141 NPSC-----RNYKMDTLEPPGDEF-----SGWACPYDAKHANVALFADGKJYSATVT 188
136 SPVCTYLNRRRRSS-----DQVFMIDSKCESKGRCSFNPNVNTVSVVINESJFSQWYI 189
189 DFLAIDAVIYRSLGESPTLRTVKHDSKWLKPEYFVQA-----VQYGD-YIYFFREIAV 241
190 DFMCTDAAIERSLT-KRMQJRTDOHNSKMLSEPFVDAHVIPDGTDPNDKAVYFFPKERLT 249
242 EYTMGKGVFPRVAQVCKXDMGGSQVLEKQWTSFLKARLNCSPVGDG--HFVFNILQAV 299
250 DNNRSTKQIHSMIARICPNIDTG-QRSLVNKMTFLKARLVCSVTDEDPETHFDELEV 308
300 TDVIRINGR-DVYVATFSPYNSIPSSAVCAVDMLDIASVFTGRFKECKSPDSTWTPVD 358
309 FLVETDNPRTLVYGIPT-SSSVFKGSAVCVYHLSDIQT-VFNGPFAHKEGPNHQLISY-Q 367
359 ERVFKPRPGCCAGSSSLERYATSNFFPDTLNFIKTHLPXDEAVPSIPNRPWFLRTMVRY 418
```

```

D 368 GRTPYPAFGTCPOGAFTPNVTKTQFDDVWVTFIRNHPPLMYNSISPIHRRPILVIGTGY 427
Y 419 RLTKIAVCTAAGPYQNHVTVFLGSEKGIILKFLARLNSOFKNSDPLFEMSVMSEKKS 478
D 428 KYKIADVDRUNADGRVHVFLGTDRCTQKVVVLTNS-ASGEILLEELVFXNH--- 483
Y 479 YDGVEDKRLVGMOLDRASSSLVAFSTCVIKVPLGCRHKKYKTCIASRDPYCGWIKK 538
D 484 -----VFITINISKKQGLIYSSNFGVSVSLHRCCHYGTACACCCCLARDPYCAM--D 535
Y 539 GGAGSHLSPNSLTPFQDIERGHTDGLGDCHNSFVALAGHSSSLPSTTTSDSACE 595
D 536 GHSCSRPTPKGRARRAQVHGN? LPOCRFENKATFNAAELVQYGVNNSPLF 591

ESUL 15
66498
-sma F protein precursor - mouse
/Species: Mus musculus (house mouse)
/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #next_change 24-Nov-1999
/Accession: S66498
/Inagaki, S.; Furuyama, T.; Iwahashi, Y.
EBS Lett. 370, 269-272, 1995
/Title: Identification of a member of mouse semaphorin family.
/Reference number: S66498; MUID: 95384603; PMID: 7456901
/Accession: S66498
/Status: preliminary
/Molecule type: mRNA
/Residues: 1-834 <INA>
/Cross-references: EMBL:S70463; NID:G1110639; PDB:1AB1B4.1; PDB:G1110699
/Superfamily: semaphorin
/1-21/Domain: signal sequence #status predicted <SIG>
/12-834/Product: N-glycophorin #status predicted <NAT>

Query Match 12.78; Score 692; DB 2; Length 834;
Best Local Similarity 26.0%; Pred. No. 3.5e-38;
Matches 232; Conservative 131; Mismatches 126; Indels 204; Gaps 34;

/ 65 NGILYIARSHIYTVDTSDTSHETVCSKKLTWKSQADYDTCRMKKG-HKDBCHNPIKY 123
D 62 SGLLYGAREALFAFSV-----EALQLQSAISWEAPAEKKECTQKGNQTECFNFIFF 116
Y 124 LLKKNDDALFVCGTNAGNPNSCRNYKMDTEPPGDEF-SGVARCEYDAKHANVALPADGKL 182
D 117 LOPNSSHLVCCGYAQPKCTYINMLPTLDRABEDGKCKCPYDPAKHTGLLVGEL 176
Y 183 YSATVDFLAIDAVIYRSLSGESPTLRVTKSDSKMLKEPYVQA-----VOYGDVI 232
D 177 YSATENNFLOTEPVILRYMGTHSIKT-EYLAFLAKRPFVGSAPVFSYGSFTGDDK1 235
Y 233 YEFRELAVENTYGVYPRVAVQCKNDMGSCRYLEKQTSFLAPLNCSTFGDSHEY 232
D 236 YFFSERAVEYDYSQVAVARVAVCKGMGGA-RLQKKAITFLKARLVCSAP-ENKYY 293
Y 293 FNILQAVTDMIRINGROW-LATFSTPYNSIPGAVCAVDMLDIASVFTSPFKEOKSPS 351
D 294 FNQLKAVHLRAGSWNTTFGVQASWGDYLSAVCEYQLECIQQVFSEPYKEYSEQAQ 353
Y 352 TWTVFDERVPEKRPCCGACSSSEIER-YATSNRPFDQTNFKTNFLNDEAVPSIFRFA 410
D 354 KWARYTDP-VFSFPGSGCINNWHFNDNYTSSLAFDNTLAFIKHFUMEDQVAF-LREL 412
Y 411 PLRTWAVRLTKIAVDTAAG-PVGNHTVWFLGAEKGIILKFLARLNSOGLNDSE--- 464
D 413 LVKXNTNF--THVTADVPLDGLGYTVLITGTCGQNLKAV-----SLGPNWH 459
Y 465 PLBEMSVNSEKSYDGVDEKXKINGMO-DRASSSLVAVPTQVYKVLPCRCEPHKCKKI 524
D 460 MWRELQVDEEP-----VFSLVLSQSKXVIFAGSRSLQVLSLADCTWYFEC-VQ 564
Y 525 CIASRDPYCGWKEGGAC-----SHLS--FNSRLTFEGDIER----- 569
D 509 CVLARDPYCAMNVTNRCVATTSCRSQSFVGVHAKLDTSKXCNQVYIKVRSIPKNTY 568

```

Search completed: October 23, 2003, 17:11:31  
 Job time : 30 secs

```

QY 560 -GNTDGLGDCHNSFVALNGH---SSLLPSTTTSDTAQEGYESRQGMLOWKHLLDSPS 615
D 569 VSGTDLVLPCHLSSNLAFAHWTFSQDLAEQPGSFYDTGLQALVW----- 616
Y 616 TDFLGAUSSHN-----HQDKKG--VIRSSYLKG-----HDCLVPTVTLIAIVILAFV 660
D 617 ----AAQSRHSGPYFCYSEBQGTFLAAESTYLVAWAGSSVTLARAPLENLGLWKLAVYA 672
Y 661 MGAVFSSGTYVVCVDHRRKQVAVVQKEXELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 720
D 673 LGAVCL-VILLVLVLSLRR-----LRSELE-----KGAKAS-----BR 704
Y 721 ALITPLVINGKLATPCTAKMGLKADQHHLULALPTPSTPTLQOKRKPSPGSRKN--- 777
D 705 TLVVPD-----ELPKEPASHPF--RPGPETDEKLMDPV 735
Y 778 -----ERNQNLINACTKMDPPYGVSPVPTDLPRLASPSHIFSVWVPTTCQGYCHEYVDQ 832
D 736 GYYVSDGSLKIVPGHARCQSGGPPSPPPQIPGCPUPS--PTRLHLLGGSNNSNANGYVR- 792
Y 833 FKMSEVAQMALEDQAATLEYKTIKEHLSSKSPNHYNLVLENLDSLEPKYVPORE 585
D 793 -----LQJGSEDRGGS-----GRPLFELADELRPKLQQRQ 822

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model.

run on: October 23, 2003, 17:09:07 : Search time 17 seconds  
(without alignments)  
2349.285 Million cell updates/sec

title: US-09-856-681-2  
effect score: 5450  
sequence: 1 MRSKALLVFTLLHPAGAGF.....PRKPSFAPLS\*SMKENDACT 1030

scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 127863 seqs, 47026705 residues

total number of hits satisfying chosen parameters: 127863

minimum DB seq length: 0  
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%  
Listing first 45 summaries

SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB	ID	Description
1	5450	100.0	1030	1	SM6A_HUMAN	Q9H2E6 homo sapien
2	4346.5	79.8	888	1	SM6A_MOUSE	Q35464 mus musculus
3	2013.5	36.9	988	1	SM6B_HUMAN	Q9H2E3 homo sapien
4	1937	36.6	657	1	SM6B_RAT	Q70142 rattus norv
5	1935	36.6	886	1	SM6B_MOUSE	Q54851 mus musculus
6	1523.5	27.8	930	1	Q9H2E2	Q9H2E2 homo sapien
7	1493.5	27.4	931	1	Q9H2E3	Q9H2E3 mus musculus
8	1431	27.4	960	1	SM6C_MOUSE	Q9H2E3 rattus norv
9	971.5	17.8	730	1	SM6A_SCHAM	Q25473 schistosom
10	932.5	17.1	771	1	SM6A_MOUSE	Q24322 drosophila
11	859	15.7	772	1	SM6A_RAT	Q63448 rattus norv
12	853	15.7	772	1	SM6A_MOUSE	Q63448 mus musculus
13	852.5	15.6	771	1	SM6A_MOUSE	Q14563 homo sapien
14	845	15.5	712	1	SM6A_MOUSE	Q24972 tribolium c
15	845	15.5	778	1	SM6A_MOUSE	Q9H2E2 brachydanio
16	845	15.5	860	1	SM6A_MOUSE	Q9H2E1 brachydanio
17	839.5	15.4	772	1	SM6A_MOUSE	Q9H2E0 gallus gal
18	830	15.2	1374	1	SM6A_HUMAN	Q13591 homo sapien
19	826	15.2	749	1	SM6B_HUMAN	Q13214 homo sapien
20	825.5	15.1	1077	1	SM6A_MOUSE	Q62217 mus musculus
21	803	14.7	748	1	SM6B_MOUSE	Q62177 mus musculus
22	802	14.7	764	1	SM6B_MOUSE	Q9H2E6 brachydanio
23	797	14.6	777	1	SM6D_HUMAN	Q9H2E6 homo sapien
24	793.5	14.6	761	1	SM6C_MOUSE	Q9H2E6 gallus gal
25	789	14.5	712	1	SM6A_MOUSE	Q9H2E6 homo sapien
26	766	14.1	1053	1	SM6B_MOUSE	Q63519 mus musculus
27	762.5	14.0	705	1	SM6A_MOUSE	Q24323 drosophila
28	762	14.0	775	1	SM6B_MOUSE	Q15041 homo sapien
29	752	13.8	751	1	SM6C_MOUSE	Q42236 gallus gal
30	751	13.8	561	1	SM6C_MOUSE	Q9H2E6 mus musculus
31	750.5	13.8	697	1	SM6A_MOUSE	Q9H2E6 schistosom
32	749	13.7	785	1	SM6B_HUMAN	Q13275 homo sapien
33	744.5	13.7	785	1	SM6F_MOUSE	Q88632 mus musculus

RESULT :  
SM6A\_HUMAN  
ID SM6A\_HUMAN STANDARD; PR7: 1030 AA.  
AC Q9H2E6; Q9P2H9;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1)  
DE (SEM6A-1)  
GS SEM6A CR K2AAL368.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND INTERACTION WITH EVL.  
RA MEDLINE=20564339; PubMed=1093894;  
RA Klostermann A., Lutz B., Gertler F., Behl C.;  
RT "The orthologous human and murine semaphorin 6A-1 proteins  
RT (SEM6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated  
RT phosphoprotein-like protein (EVL) via a novel carboxyl-terminal  
RT zyxin-like domain.";  
RL C. Biol. Chem. 275:39647-39653(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
PC TISSUE=Brain;  
RX MEDLINE=20191126; PubMed=10718195;  
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Chazara C.;  
RT "Prediction of the coding sequences of unidentified human genes. XVI.  
RT the complete sequences of 150 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:165-73(2000).  
CC -!- FUNCTION: Can act as repulsive axon guidance cues. May play a role  
CC in channeling sympathetic axons into the sympathetic chains and  
CC controlling the temporal sequence of sympathetic target  
CC innervation (By similarity).  
CC -!- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9H2E6-1; Sequence=D:played;  
CC Name=2;  
CC IsoId=Q9H2E6-2; Sequence=VSP\_007113;  
CC Note=No experimental confirmation available;  
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC -!- SIMILARITY: Contains 1 Sema domain.

ALIGNMENTS

34	739	13.6	775	1	SM3E_MOUSE	P02275 mus musculus
35	737	13.5	751	1	SM3C_MOUSE	Q62181 mus musculus
36	735.5	13.5	862	1	SM4D_HUMAN	Q92854 homo sapien
37	733.5	13.5	785	1	SM3E_CHICK	Q42237 gallus gall
38	733	13.4	751	1	SM3C_HUMAN	Q99985 homo sapien
39	732.5	12.9	766	1	SM27_BRAVE	Q9YH44 brachydanio
40	692	12.7	834	1	SM4C_MOUSE	Q64151 mus musculus
41	681	12.5	832	1	SM4B_HUMAN	Q9H282 homo sapien
42	672	12.3	761	1	SM4A_HUMAN	Q9H381 homo sapien
43	664	12.2	776	1	SM4F_RAT	Q92143 rattus norv
44	658	12.1	777	1	SM4F_MOUSE	Q92123 mus musculus
45	656	12.0	782	1	SM4B_MOUSE	Q62179 mus musculus

```

DR BMSJ: AF279556; AAG23378.1;
DR EMBL: AB037789; BAA32606.1; ALT_INIT.
DR GenBank: HGNC:1073; SEMA6A.
DR WIM: 605985;
DR InterPro: IPR003459; Plectin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema.
DR SMART: SMO0423; PSI: 1.
DR SMART: SMO0633; Sema.
DR Signal: Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
DR Developmental protein; Alternative splicing.
DR SIGNAL 1 19
DR CHAIN 19 1017
DR DOMAIN 19 649
DR TRANSMEM 650 673
DR DOMAIN 671 1033
DR DOMAIN 54 451
DR DOMAIN 232 819
DR CARBOHYD 33 31
DR CARBOHYD 49 49
DR CARBOHYD 65 65
DR CARBOHYD 282 282
DR CARBOHYD 434 434
DR CARBOHYD 461 461
DR VARSPLIC 576 576
DR SEQUENCE 1030 AA; 124369 MW; A57B75C0A0BC4E34 CRC644;
Query Match
Best Local Similarity 100.0%; Score 5450; DE 1; Length 1030;
Matches 1030; Conservative 0; N:smatches 0; Indels 0; Gaps 0;
1 MRSEALLVFTLLHFAAGAFDESPIS:SGHNTYKQVYVGHKPGKNTTORHPLDIOM 60
1 MRSEALLVFTLLHFAAGAFDESPIS:SGHNTYKQVYVGHKPGKNTTORHPLDIOM 60
61 IMKMGTLNIAARDHYTVDTSTETELYSKSLTWKSRQADYTCRYKKHKBOSCHNF 120
61 IMKMGTLNIAARDHYTVDTSTETELYSKSLTWKSRQADYTCRYKKHKBOSCHNF 120
121 IKVLLKNDALFVGTNAFNSTCRNYKMTLEPPDFEFGKACQPYDAKHANVALPADG 180
121 IKVLLKNDALFVGTNAFNSTCRNYKMTLEPPDFEFGKACQPYDAKHANVALPADG 180
181 KYSAIVTDFDAIDAVIVRSGLSESTLSTWYKSDSKMLREYFVQVNDYGVIVFEFEIA 240
181 KYSAIVTDFDAIDAVIVRSGLSESTLSTWYKSDSKMLREYFVQVNDYGVIVFEFEIA 240
241 VEYNTNGKVYVPRVAVQVCKNKGSSQVLEKQNTSFUKARLNCVPGDSHFYFNLCQVT 300
241 VEYNTNGKVYVPRVAVQVCKNKGSSQVLEKQNTSFUKARLNCVPGDSHFYFNLCQVT 300
301 DVIRINGRWVATSTPNSIPSSAVAYMDLQASHTGREFEQQKSPDSTWTFVDEP 360
301 DVIRINGRWVATSTPNSIPSSAVAYMDLQASHTGREFEQQKSPDSTWTFVDEP 360
361 VPKPRGCCAGSSSLERVATSEFFDQTNFLKTHPLMDEAVPSIFNRFNLELWVRYRL 420
361 VPKPRGCCAGSSSLERVATSEFFDQTNFLKTHPLMDEAVPSIFNRFNLELWVRYRL 420
421 TKIAYTAAGPVNHVTFVLSGKERTLKFARIGNSHFLSDSIFLEMSVYNSEKCSYD 480
421 TKIAYTAAGPVNHVTFVLSGKERTLKFARIGNSHFLSDSIFLEMSVYNSEKCSYD 480
481 GVEKDKIMGMCLDRASSSVYAFSTGVIVKVPVGRERKCKYKTIAGSDYQCKWIKSGG 540
481 GVEKDKIMGMCLDRASSSVYAFSTGVIVKVPVGRERKCKYKTIAGSDYQCKWIKSGG 540
541 ACSHLSPNSRLTFEQDIERGNTDGLGDCCHNSFVALNGHSSALLPSTTSDTAGGVESR 600
541 ACSHLSPNSRLTFEQDIERGNTDGLGDCCHNSFVALNGHSSALLPSTTSDTAGGVESR 600
601 GGMLDKWLKLDSDSDTDLPLGKNSRNHQQKGVIRESTLKHQCQLVPTLLAIAVILAFV 660

```

```

661 GGMLDKWLKLDSDSDTDLPLGKNSRNHQQKGVIRESTLKHQCQLVPTLLAIAVILAFV 660
661 XGAVFSGITVYCHCHARKKDYAVYQKKEKELTHSRGCSKSSVTKLSGLTQOTOSKDEKPE 720
661 NGAVFSGITVYCVCHCHARKKDYAVYQKKEKELTHSRGCSKSSVTKLSGLTQOTOSKDEKPE 720
721 AILPLVXHGKGLATPGNTAKMLKADQHLDELTALETPPESTPTLQCKKPSRSGSRWERN 780
721 AILPLVXHGKGLATPGNTAKMLKADQHLDELTALETPPESTPTLQCKKPSRSGSRWERN 780
781 QYLINACTKOMPNGSPVITPLPLRASPSPHISVWVLPITCCGYOHEHYVQCPKSEVAC 840
781 QYLINACTKOMPNGSPVITPLPLRASPSPHISVWVLPITCCGYOHEHYVQCPKSEVAC 840
841 MALEQDAATLEVKTKKEHLSKSPNHQNLVEMLDSPPKVPQREASLGPFGASLSQTLG 900
841 MALEQDAATLEVKTKKEHLSKSPNHQNLVEMLDSPPKVPQREASLGPFGASLSQTLG 900
901 SKRLEMHSSSYGVYKRSYPTSLTSHQATTLKRNNTNSSSHLSRNQSGFSGENPP 960
901 SKRLEMHSSSYGVYKRSYPTSLTSHQATTLKRNNTNSSSHLSRNQSGFSGENPP 960
961 PAPQPVDSIQVSSSCPSGQAVTVSPQSLNAYNSLTPSLGKRTPSLKPQVPPKPSFAPLS 1020
961 PAPQPVDSIQVSSSCPSGQAVTVSPQSLNAYNSLTPSLGKRTPSLKPQVPPKPSFAPLS 1020
1021 TSMKPNDACT 1030
1021 TSMKPNDACT 1030

```

## RESULT 2

```

SEMA_VQJUSE
ID SEMA_VQJUSE STANDARD; PRT; 889 AA.
AC Q35464;
CT 30-MAY-2003 (Rel. 39, Created)
CT 30-MAY-2003 (Rel. 39, Last sequence update)
CT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1)
DE (SEMA6A-1) (Semaphorin Q) (Sema Q).
GN SEMA6A CR SEVAQ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
PN [1]
PP SEQUENCE FROM N.A.
RX MEDLINE=97348468; PubMed=9204478;
RA Zhou L., White F.A., Lentz S.L., Wright D.E., Fisher D.A.,
RA Snider W.D.;
RT "Cloning and expression of a novel murine semaphorin with structural
RT similarity to insect semaphorin I."
RL Mol. Cell. Neurosci. 9:26-41(1997).
RN [2]
RP INTERACTION WITH EVL.
RX MEDLINE=23564339; PubMed=10593894;
RA Klottermann A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins
RT (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated
RT phosphoprotein-like protein (EVL) via a novel carboxyl-terminal
RT zyxin-like domain."
FL J. Biol. Chem. 275:39647-39653(2000).
CC [1-] FUNCTION: Can act as repulsive axon guidance cues. May play a role
CC in channeling sympathetic axons into the sympathetic chains and
CC controlling the temporal sequence of sympathetic target
CC innervation.
CC [2-] SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.
CC [3-] SUBCELLULAR LOCATION: Type 1 membrane protein.
CC [4-] TISSUE SPECIFICITY: PARTICULARLY HIGH LEVELS IN SPINAL CORD,
CC CEREBELLUM, METENCEPHALON, SUPERIOR AND INFERIOR COLLICULUS,
CC Diencephalon, Olfactory bulb, and eye.
CC [5-] DEVELOPMENTAL STAGE: TEMPORALLY AND SPATIALLY REGULATED DURING

```



DEVELOPMENT.

1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

2- SIMILARITY: Contains 1 Sema domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF030430; A886408.1; -

MGD; MG1123727; Sema6a.

InterPro; IPR003652; Plexin-like

InterPro; IPR003652; Sema.

Pfam; PF01403; Sema; 1.

SMART; SM00423; PSI; 1.

SWART; SMC0610; Sema; 1.

Signal; transmembrane; Multigene family; Neurogenesis; Glycoprotein; Developmental protein.

1 15 POTENTIAL.

2 19 888 SEMAPHORIN 6A.

3 19 649 EXTRACELLULAR (POTENTIAL).

4 650 670 POTENTIAL.

5 671 888 CYTOPLASMIC (POTENTIAL).

6 56 491 SEMA.

7 792 819 PRO-RICH.

8 33 33 N-LINKED (GLCNAC). (POTENTIAL).

9 49 49 N-LINKED (GLCNAC). (POTENTIAL).

10 65 65 N-LINKED (GLCNAC). (POTENTIAL).

11 282 282 N-LINKED (GLCNAC). (POTENTIAL).

12 434 434 N-LINKED (GLCNAC). (POTENTIAL).

13 461 461 N-LINKED (GLCNAC). (POTENTIAL).

14 898 AA; 96075 MW; C1C94FC2EB345C57 CRO44.

Query Match 79.8%; Score 4346.5; DB 1; Length 888.

Best Local Similarity 94.4%; Pred. No. 3-5e-272;

Matches 916; Conservative 24; Mismatches 23; Indels 1; Gaps 1;

1 1 MRSEALLYTLHIFAGAGPEDEPISISHGNVTKCYPVVGHKGRNTTQRR:DIOM 60

2 1 MRPAMLSCTLLICAGAGPEDEPISISHGNVTKCYPVVGHKGRNTTQRR:DIOM 60

3 61 IYINGNGLYANRQHTVTVDTSHBEIYCSKLLNKEBCQAVYCRNKGKQKORCHE 120

4 61 IYMNRTLYVAARDHTVTVDTSHBEIYCSKLLNKEBCQAVYCRNKGKQKORCHE 120

5 121 IKVLLKNDALFVCGTNAPNRCNRYKNTLEPFGDEFSGMACQPYDAKHANVAFAG 180

6 121 IKVLLKNDCTLVCGTNAPNRCNRYKNTLEPFGDEFSGMACQPYDAKHANVAFAG 180

7 191 KYSATVTDFAIDAVYRSLGSEPTRTYKHSKMLKEEYFYQAVDYDYIYFFERE:LA 240

8 181 KYSATVTDFAIDAVYRSLGSEPTRTYKHSKMLKEEYFYQAVDYDYIYFFERE:LA 240

9 241 VEYTMGVVPEPVACQKNDYGSQRYLQKQTSRLKARLNCVPGDSHFYENLQAVT 300

10 241 VEYTMGVVPEPVACQKNDYGSQRYLQKQTSRLKARLNCVPGDSHFYENLQAVT 300

11 301 DVIRNGRDVJATFTSTPYNSIPGSAYCAYMDLIDIASVFTSRFEKQSPSTWTPVEDER 360

12 301 DVIRNGRDVJATFTSTPYNSIPGSAYCAYMDLIDIANVFTSRFEKQSPSTWTPVEDER 360

13 361 VPKRPCCAGSSSLERYATSNRPDITLPIKTHPNDVDAVSIKRPWFLRTWRYLJ 420

14 361 VPKRPCCAGSSSLERYATSNRPDITLPIKTHPNDVDAVSIKRPWFLRTWRYLJ 420

15 421 TKIADVTAAAGYQKHTVYVLGSEGGILKEARTGNSGFLNCSLFLPEEMSVYSEKSYD 480

16 421 TRIADVNAAGYQKHTVYVLGSEGGILKEARTGNSGFLNCSLFLPEEMSVYSEKSYD 480

17 461 GVEDKRIMGOLDRASSSLYVAFSTCVIKVPLGRCHRGKCKTC:ASRDPYCGW:KEGG 540

Db 481 GVEDKRIMGOLDRASSSLYVAFSTCVIKVPLGRCHRGKCKTC:IASRDPYCGW:KEGG 540

Qy 541 ACSHLSPNSRLTTEQDIERGNIDGLGCHNSFVALNGHSSLSLPSTTTSDS:TAQGYEFSR 600

Db 541 SCAHLSPLSLTTEQDIERGNIDGLGCHNSFVALNGHASSLYVSPSTTTSDS:ASRSGYESR 600

Qy 601 GGLMDKHLDDSPSTDFPLGAVSSHHNQKKGKGVIRESYLKGHCQLVPVTLAIAVILAFV 660

Db 601 GGLMDKHLDDSPSTDFPLGAVSSHHNQKKGKGVIRESYLKGHCQLVPVTLAIAVILAFV 660

Qy 661 XGAVESGITVYCVCDHRRKQDAVYQKKEKELTHSRGSMSSVTKLSGLPGTCSKDPKEE 720

Db 661 XGAVESGITVYCVCDHRRKQDAVYQKKEKELTHSRGSMSSVTKLSGLPGTCSKDPKEE 720

Qy 721 ALITPLMHNGKLTATPNTAKMLIKADQHLDLTALPTPESTTTLOCKRXPGRSGREWEHRN 780

Db 721 ALITPLMHNGKLTATPNTAKMLIKADQHLDLTALPTPESTTTLOCKRXPGRSGREWEHRN 780

Qy 781 QNINACTKDMPPGSPVPTDLPBASPSHIPSVVPLITQGGYCHYVVDQPKMSE:VA 839

Db 781 QNINACTKDMPPGSPVPTDLPBASPSHIPSVVPLITQGGYCHYVVDQPKMSE:VA 840

Qy 840 QMALEDQATLEVKTIKEHLSSKS 863

Db 840 QMALEDQATLEVKTIKEHLSSKS 864

RESULT 3

SM6B HUMAN STANDARD; PRT; 888 AA.

10 Q9H3T3; Q9NRK9;

16 OCT-2001; Rel. 42; Created:

16 OCT-2001; Rel. 40; Last sequence update:

15-SEP-2001; Rel. 42; Last annotation update:

DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).

GN SEMA6B CR SEMAZ.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A. (ISOFORM 1).

PC TISSUE=Brain;

RA Kimura T., Ishida H.;

RI Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN 121

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE=21248680; PubMed=11350127;

RA Cortea R.G., Sasahara P.Y., Bengtsson M.H., Karayama M.L.H.,

RA Salim A.C.M., Brentani V.M., Sogayar M.C., de Souza S.J.,

RA Simpson A.J.G.;

RA "Human semaphorin 6b".

RI Genomics 73:343-348(2001).

FD 1- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS

CC 1- SYSTEM DEVELOPMENT (BY SIMILARITY).

CC 1- SUBCELLULAR LOCATION: Type I membrane protein.

CC 1- ALTERNATIVE PRODUCTS.

CC Event=Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=1;

CC Iso=3; Q9H3T3-1; Sequence=Displayed;

CC Name=2; Synonym=6B.1;

CC IsoId=Q9H3T3-2; Sequence=VSP\_006044, VSP\_006045;

CC 1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC 1- SIMILARITY: Contains 1 Sema domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



InterPro: IPR003659; Pfam: like.  
 InterPro: IPR001227; Pfam: like.  
 Pfam: PF01403; Pfam: like.  
 SMART: SM00423; SMART: like.  
 SMART: SM00630; SMART: like.  
 Signal: Transmembrane; Multigene family; Neurogenesis; Glycoprotein; Developmental protein.  
 T SIGNAL 26  
 CHAIN 27 887  
 DOMAIN 27 605  
 TRANSMEM 606 626  
 DOMAIN 627 597  
 DOMAIN 239 549  
 CARBOHYD 75 75  
 CARBOHYD 156 156  
 CARBOHYD 168 168  
 CARBOHYD 292 292  
 CARBOHYD 387 387  
 CARBOHYD 442 442  
 CARBOHYD 463 463  
 SEQUENCE 887 AA; 95752 MW; 05543F3F232CD3C1 CRC64;

Query Match: 36.6%; Score 1997; DB 1; Length 887;  
 Best Local Similarity 42.7%; Pred. No. 7, 1e-121;  
 Matches 433; Conservative 132; Mismatches 260; Indels 193; Gaps 25;

1 2 REALLLYTLLHFAAGGPERSEPISSHONYTKQYFVCHKKQGNNTOR--HADIQ 59  
 10 RPALLFELLLLARVTHGSEDFEPLSVAPDYLSHYFVFGSGGGRJTPAEGADLNH 69  
 60 MIMVNGTLYIAARHIYVVDIDHTEIIVCSKLTWKSROADYDTPMKKQKQECHN 119  
 70 RVLRVNTPIGCRONKYVLEPSTELKCRKLTWPSNPSDIDVPMKQKQGERH 129  
 120 FIKVLKNDQDAJFVCKTAFKPSQNYKMTLRFQDFEFSQVACRYDAKNAVAFAD 179  
 130 FUKVLLRPSTLFCVGSNAEPICANYSMOTLQLLQNIQSNWACPYDKKANVAFSD 189  
 180 GLYSATVDELDAIDAVIVESGESSTLVKHSKMLKEPYFNGAVGVYVFFPEI 239  
 190 GMLFTATVDELDAIDAVIVESGSRPTSTVXHSKQKPEPYFNAVAVGSHVYFFPEI 249  
 240 AVEYNWGVVPRVACVCKNDGQGVLEKQWTFKARLNCVSPVDSHYFVILCAV 299  
 250 ANEPVLEKVVSRVAVCNVQVSPVLEKQWTFKARLNCVSPVDSHYFVILCAV 309  
 300 TVIRNGDVLATSTYNIIFSAVCAVYKLDIASVFTGRFEQKQSPSTVTFVDE 359  
 310 TGWVSLGSRVILAVFTSNIFSAVCAVQVAVFQVAVFQVAVFQVAVFQVAVF 369  
 360 RYKPRPGCCASSSLERATSNPEEDTNEKTHPLMDEAVPSIENRPMFLRTMYR 419  
 370 QVPRPFGCAAPGM--QYNASVAFDELAFVKTFLMDEAVPSLGHSPWAVRTLRQ 427  
 420 LTKIATVTAAGPVNHTVFLGSEKILKFLAR--IGNSGELNDSFLEENSVNKSCK 477  
 428 LRVAVDVAGVNGNTVFLGSEVGVKFLVKNPASVSGTGTGSLFEEETVTPESC 487  
 478 ---SYDGVTEKRMVQVLRASSLSVAFSTVIVKVLQRCERHGKCKKTCASREPYCG 534  
 488 GRSSSAGEVQRLLSLELDAASGGLLAAPRCVAVFVAVQVLYSCMKNCVGSQCPVCG 547  
 535 NKEGACACSHLSPKSLTQDIERNTGCGDCHNSFVALKCHSSLLPSATTTEDTAQ 594  
 548 WAPD-GSCIFPREGTSATPEQVSGASTSGDGC----- 580  
 595 EGVESRGMLDQWKLHLLDSFDTLGAVSSHNQDKKGVIRESYLKGHDQVVPVTLAIA 654  
 581 -----TGJLRASLSDDRAGLVSNLLVTS 664  
 655 VILAFVNGAVFGSCTVYCVCDHRRKQDAVAVRKEKE--LTSRRGSMSSVTKL----- 705  
 605 SVAAFVVGAVSGVSGVGVGLRERRELA--RRKDEAILAHGGSEAVLSVSRJGERRGT 662

CY 706 --SGLFGTQSKQPKBEALVPLMHNGKLTATPNTAKMLKADCHHDLVLTALTPPSTPT 763  
 DB 663 GTGGRGAGGGPGFPEALAPLQNGV-----TKAALLHGPHSDSGLLTPPOTP- 715  
 CY 764 LOOKRXPSR-----GSREWRNQLINACTK-----CXPNGSPVIFTD-- 802  
 DB 716 LPQKRUPVTPHFAHALGPRANDHSHALLSASASTSLLLAHTAPEOPP-----VTESS 770  
 CY 903 -----LPLRASPSPHPSVVVLDITCGYQHYEVDPKMSVAQVALEDOAATLEYKTIK 856  
 DB 771 PESRLCAPRSCRASHPGDFLTP-----HASPDRRVVSAPTGPLDSSSVG----- 816  
 CY 857 EHSSSKSPNHGVNLENLDSLP-PKVPQCEASL-----CPCASLCTGCSKLEMHHS 909  
 DB 817 -----DCLPGWSPPTATSSURRPGHPPTAAARRT-----KT 845  
 CY 910 SSYGVQVYKRSYPTNSLTFSHQATTAKRNTNNSNSSLRNQSGRQD-NPPAP 963  
 DB 950 FNSG----EAPGGRPRRHA-----PADSTHL---LPCGIGERTAPPVP 987

RESULT 5  
 SM6B MOUSE  
 ID SM6B MOUSE STANDARD; PRT: 886 AA.  
 AC 054951;  
 DT 30-MAY-2000 (Rel. 33, Created;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update);  
 DT 16-OCT-2001 (Rel. 43, Last annotation update);  
 DE Semaphorin 6B precursor (Semaphorin V-B) (Sema VIB) (Semaphorin N);  
 DE (Sema N).  
 GN SEVA6B OR SEMAN  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 PP SEQUENCE FROM N.A.  
 EX MEDLINE=98027184; PubMed=9361278;  
 RA Eckhardt F., Behar O., Calautti E., Yonezawa K., Nishimoto I.,  
 RA Fishman M.C.;  
 RL Mol. Cell. Neurosci. 9:439-419 (1997);  
 CC -- SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE  
 CC PROTOONCOGENE C-SRC.  
 CC -- SUBCELLULAR LOCATION: Type: membrane protein.  
 CC -- TISSUE SPECIFICITY: DURING DEVELOPMENT IT IS EXPRESSED IN  
 CC SUBREGIONS OF THE NERVOUS SYSTEM AND IS PARTICULARLY PROMINENT IN  
 CC MUSCLE. IN ADULTHOOD, IT IS EXPRESSED UBICUITOUSLY.  
 CC -- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -- SIMILARITY: CONTAINS: Sema domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EVBL; AF36585; AAC00433.1;  
 DR MGI:1202889; Sema6b.  
 DR InterPro; IPR001659; plexin-like.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF01403; Sema; 1.  
 DR SMART; SM00423; PS; 1.  
 DR SMART; SM00630; Sema; 1.  
 KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 KW Developmental protein.  
 FT SIGNAL 26  
 FT CHAIN 27 886  
 FT DOMAIN 27 605  
 FT TRANSMEM 606 626  
 FT POTENTIAL.



DOXAIN 25 604 EXTRACELLULAR (POTENTIAL).  
 TRANSXEM 605 525 POTENTIAL.  
 DOXAIN 626 230 CYTOSOLIC (POTENTIAL).  
 DOXAIN 233 343 SEMA.  
 DOXAIN 662 667 POLY-PRO.  
 DOXAIN 752 755 POLY-PRO.  
 CARBOHYD 70 70 N-LINKED (GLNAC). (POTENTIAL).  
 CARBOHYD 286 286 N-LINKED (GLNAC). (POTENTIAL).  
 CARBOHYD 437 437 N-LINKED (GLNAC). (POTENTIAL).  
 VARSPLIC 184 223 Missing (in isoform 2).  
 VARSPLIC 556 556 /FTID=VSP\_006346.  
 Y -> YV2PGRGSPGTPSPDAPHRFQSSGLVHTS  
 (in isoform 2 and isoform 3).  
 /FTID=VSP\_006347.  
 I -> V (IN REF. 2: AAL70299).  
 R -> K (IN REF. 1).  
 P -> T (IN REF. 1).  
 SEQUENCE 930 AA; 99652 NW; 8AF8814AD9C84C98 CRC64;  
 Query Match 27.8%; Score 1515.5; DB 1; Length 930;  
 Best Local Similarity 35.8%; Pred. No. 7.7e-90;  
 Matches 380; Conservative 136; Mismatches 356; Indels 189; Gaps 32;  
 6 LLYFTLHFAAGAPEDSESPIS-SHKNYTKYPPVFGKFKGRNTTCRRRLCIQXIMN 65  
 13 LLLLLSLPH-TQAAFPQDPLPLLSLQSTFLSWERGLEDDAAAE-L-GDFQRELTIN 70  
 66 GLVY-ARSDHYTYDIDT-SHT-BEYCSKKY-TWSPQADVDTCMKGKHKECHNF-KVL 124  
 71 RLLVNAHDHVFQDQAELEGVLPKXK-TWESQ--DVNCAVRGKLTDCEVYIRVL 129  
 125 LKNDLALFVGTNAFNPSCRYKVTLEPPDEPFSGVARCFYDAKHANVALFADGKLYS 184  
 129 VPWDSQTLACGTSFVSVCSYGYSTIQQGEELSGQARCFPATGCVAFREGSLYS 188  
 135 ATVDFLAIDAVIVASLESPTLR-TWKSQSKMLKEFPYQAVQVGYIYFFRDIAYEYN 244  
 189 ATADPQASCAVYVASTLQFPQPRSAKYDSKMLREPHFVQALEHCHVYFFRVSVEDA 248  
 245 TWGKTYPPRVACVCHVDGGGQORVLEKONTFLKARLNCSPVDSHFEN-LQAVTVIR 304  
 249 PLGRVCFGRVAVRCKEDYGGSPRAEDRWTSFLK-LRLNCFVEGDSFEYVQLQALTPN 308  
 305 INGVQWLATSTPNYS-PGSAVCAYMCLD-ASVFTGFRKECKSPDSTWTVVPDVPKP 364  
 309 LHGRNALGVTTQNTS-PGSAVCAYFLDE-ERGFEGKEKESRLDGWTVSDRPPSP 369  
 365 RFGCAGSSSLERYATSNFEDTLNFKTHPLMPEAYPS-FNRPWFLRTVYRVLTKIA 424  
 369 RFGSCAGVGAALFSSRDLPDQVLTFAKHFLLQFAVPPVTHQ-LITL-SRALLTQVA 427  
 425 VDTAGPYQNTWTF-LGSEKGLILKFLARIGNSGF-NDSLFLBENSUVNSKCSYGVGD 484  
 426 VDGXAGPSHNTVWFLGSDGTVLVLPDPPGRSSG-PEPILLESIDAYSPARCSKTAQ 486  
 485 --KPYMGHQLPRASSSVYVASTCVIKVPLGRCHRGKCKTKCLASROPYCGWIKERGA 542  
 487 TARR--GLLEDTGHRLLFVANSGCCVYIPLSARAHGACQPSCLASQCPYCGWSSRG 545  
 543 SHLSFNSRLTRFQDIERGNTDQ--GDCHNSFVALNGHSSLLRSTTSDTAQGYESR 600  
 546 VDIRSGGTVDQ--ACNQSSMEHGCDQDG-----ATGQSGPGGSAY--- 586  
 601 GGLMDKWKLLDSFDSTDP-LGAVSNHSHQKGVIVRESYKLGHDQVFTVLL-AIVAVLAF 660  
 587 -----GVRDLPPASASRSVPIPLLLASVAAAPA 615  
 661 MCAVFGITVYVCVDRHKKVAVVCRKEKELTHSPRG-----SKSSVTKLSGL 708  
 616 LGASVSGLLVSCAC--RR-----AHRRRGKOITPGSLRPLSLRSJARLHG- 659  
 709 FGDTSKQPKP--EAILTPLMHNGKLATPGNTAKMLKADCHHLDLTALP-PESTFTLQJ 766

DB 660 -GGPEPPEBSKGDNAVQTEGLYTFLLPPEEGYPP-----ELACLTPPSTHELTPV 709  
 QY 767 KRKPSRGSSEKERNONLINACTHKMPNMSFVPTDPLRASPSH-----IPSVVLPIT 821  
 DB 710 KHLRAAGD-PWEYNORRNA-----KEGPGRSRGGHAGGPAPRVLYRPP-- 752  
 QY 822 QCGYQHEVYDQPMKSEVAOMALEQAAATLEYKTIKEHL-----SSKSPNKGVLNVLDS- 876  
 DB 753 -----PPGCGGCG-----AVEVITLELLYLHGPQPRKGAEPAPLTSR 793  
 QY 877 -LPKVPCEASLGPFGASLSQTLGSKRLEMHSSSYGVYKESYPTNSUTRSHQATLK 935  
 DB 794 ALPPE--PAPALGGSPRPHECASPLRLDV-----PPEGRCASAPA---- 833  
 QY 936 RNNTNSNSSHL-----SRNQSFGRGNGPPAPORVDSIQVHSSQFSGQAVTVSRQPS- 988  
 DB 834 -RPALGAPAPRLGVGGGRALPFGSHRAPFALLTRV-----PSGSPRSYSGGPGKHL 883  
 QY 989 --LNAYNSLTRSGLKRTPLSKPDVPPKPSF-APJSTSMKPN 1026  
 DB 884 -YLGRPEGVGRALKRVDVEKPOLSLKPLVLGPPSSRCQAVPN 924  
 RESULT 7  
 SM6C\_MOUSE STANDARD; FRT; 931 AA.  
 ID\_SM6C\_MOUSE QWTV3;  
 DC 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).  
 GN SEMA6C OR SEMAY.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RE SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=9916821; PubMed=1049525;  
 RA Kikuchi K., Chedotal A., Hanafusa H., Ujima Y., de Castro F.,  
 PA Goodman C.S., Kimura T.;  
 PT Cloning and characterization of a novel class VI semaphorin,  
 PT semaphorin Y.";  
 PL Mol. Cell. Neurosci. 13:9-23(1999).  
 CC -- FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION  
 CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER  
 CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF  
 CC NEURONAL CONNECTIONS (BY SIMILARITY).  
 CC -- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -- SIMILARITY: Contains 1 Sema domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: A5013729; BAA76294.1;  
 DR MGD: MG21338032; Sema6c.  
 DR InterPro: IPR001627; Sema.  
 DR Pfam: PF01483; Sema; 1.  
 DR SMART: SM00630; Sema; 1.  
 KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 KW Developmental protein.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 931 SEMAPHORIN 6C.  
 FT DOMAIN 26 605 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 606 626 POTENTIAL.  
 FT DOMAIN 627 931 CYTOSOLIC (POTENTIAL).

```

T DOXAIN 234 541 SEXA.
T DOMAIN 663 669 POLY-PRC.
T DOMAIN 754 757 POLY-PRC.
T CARBOHYD 7 7 N-LINKED (GLCNAC) (POTENTIAL)
T CARBOHYD 287 287 N-LINKED (GLCNAC) (POTENTIAL)
T CARBOHYD 438 438 N-LINKED (GLCNAC) (POTENTIAL)
C SEQUENCE 931 AA: 99537 MW: 51299DS94205F125 CR564;

Query Match 27.41; Score 1493.5; CB 1; Length 931;
Best Local Similarity 35.73; Pred. No. 2e 88;
Matches 374; Conservative 145; Mismatches 153; Indels 177; Gaps 32;

Y 6 LLLVFTLLHFRAGAGFPEDSPIS:SHQNTYKCTVYVGHKPKPHNTTORHRLDQIMKIM 65
b 13 LLLSSLLPQAQAAPQPTPLG:LSLQQAASPSWFRGLEDDAAVAEL-GLSPQRFPLM 71
Y 66 GTVYIAARQHYTVDTTSHT-ELHYGSKLTKWSQAGVDTCTRMKQKIOCHNFIK 124
b 72 RLIIAANDRVFSDJCAQCEQVLPKRLTRSO--DMKCNVRSKLTDCYNYIRV 129
Y 125 LKXEDALFVGGTNAIPSCNYMOTLHPGEPFGACQTYDAKIANVAFLACKLYS 184
b 130 VPMKSQTLACGTSNPSMCKSYGITSLOQGEELSQARCFEATQSTVA:FAEGSLYS 189
Y 185 ATVTDFALDAVIVESI:GESPTLTWKHSKW:KEPYFCAKCYGVYVFESEIAVEN 244
b 190 AIAADFOASCAVYHSLGPOFFLNSAKYDSKWRBPHFVTALEHGHVYFFREKSVRA 249
Y 245 TMGVVFFRYAQNCKNGASQVLEKQCTSLKARLNFFVFDSSHYPNLIQANTDVF 304
b 250 RLGRVQFSVARVCKXNGSGSPALCBHW:SELKRLNCVSPGDS:TFEDVLQSLGPN 309
Y 305 INGRVVAITSPYNS:FGSANCAYMDL:IAVFTGRKEQKSPSTWTPDRAVNP 364
b 310 LHGRSALEFVFTQTSNPSAVCAFYDTERGFEQKPEQSLSGCAKTPVSEKVPSP 369
Y 365 RPGCAGSSILERVATSENPDDT:LNFKTHPP:MDEAVPSIKRPFRTMVPYRTKIA 424
b 370 RPSGAGVGAASSESSQDLPDVLFIKHP:LDPAVPATHQP-LLTLSPALLTQVA 429
Y 425 VDTAAGPYONHTVFLASEGK:IIKFAPIGNSGFLNLSLFJEMSVYNSKSGVDGVED 484
b 429 VDGVAQPHRNTTVLFLSGNDGTVLKVLPP-GGQSLGSEPTVLEIDVSHARCS--GRKS 485
Y 485 ----KRIKGVCLDASSLYVARS:TCV:KVP:GRCEPHGKATCTIADSPVCGM:KEGQ 542
b 486 PRAARRIGLELDTGHRLLVAFEGCTVYVLSGRCAHGAQCSGLASGSPYCGMHSRG 545
Y 541 ACSHLSFNSRLTFEQDIERGNTGCGCHRSFVALAGHSSLLPSTTSDSTAQEGYE-S 599
b 546 CMSIRGP-----GGTD-----VDLIGNQES-----TENGCQDQATGS 578
Y 600 RSGVLDKXHLDPDST:DP:GAVSENHQKGVIRBSY:KGDQVLPVTLIA:AVILAF 659
b 579 QSGFGDSAY-----GVRDLSFASASRSLP:PLSLACVAAAF 615
Y 660 VYGVAFSGITVYVCDHRRKXAVAVORKEKELTHSRGGSVSVTKLSGFGDTQSKDPRP 719
b 616 ALGASVGLLVSCAC--RRAN----RRSKDTEPG:FRPL:RS:ARLHGSPPEPPPP 669
Y 720 ---EAILTFLMNGKLTATPGNTAMKLINADQHLDTALTPTSTPTLOQRKPSGRSE 776
b 670 KDGAAQTEQVITFLPRPGSGSP-----ELACLTPTPTTTELFVXK:RASGG-P 719
Y 777 KERNQNLINACT-KDYMNGSGFVPTDLP-----LEASFSHIPSVVWV:FIHQ---GTCV-- 827
b 720 KEMNQNGNAGESGFRPPHSGSGGAPRVLPVAPPPPGCGQAVEVTTUEBLRYLRSP 779
b 828 -----EYVQPKMSEVACMALEDQAATLEVKTIKELSSKGFPHGVNLVEML 874
b 780 QPPRKGSPLASAFPTSPRPASERGASLFVD-----SSPMPGCVPPD-PL 824
b 875 DLSLFPK-----VFQREAS:GPPGASLSQTGLSKRLMEHMISSYGVYKRSYPTNLSLTSHQ 910

```

```

RESULT 5
SM6C_RAT
18 SE6C_RAT STANDARD; FRT; 96C AA.
AC QWNTJ3: QWNTM6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
GN SEMA6C OR SEMAY.
CS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
RN 1.
SEQUENCE FROM N.A. (ISCPORVS SEMA Y-L AND SEMA Y-S).
RP STRAIN=Sprague-Dawley; TISSUE=Muscle;
RX MEDLINE=9316082; PubMed=1049528;
RA Kikuchi K, Chedotal A, Hanafusa R, Ujinasa Y, de Castro F,
RA Goodman C.S., Kimura T.
RA "Cloning and characterization of a novel class VI semaphorin,
RA semaphorin Y."
RC Mol. Cell. Neurosci. 13:9-23(1999).
CC 1. FUNCTION: SHOWS GROWTH CONE COLLAPSING ACTIVITY ON DORSAL ROOT
CC GANGLION (DRG) NEURONS IN VITRO. MAY BE A STOP SIGNAL FOR THE DRG
CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
CC NEURONAL CONNECTIONS. Type I membrane protein.
CC 2. SUBCELLULAR LOCATION: Type I membrane protein.
CC 3. ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Name= isoforms=2;
CC Name=Sema Y-L;
CC IsoId=O9WLT3-1; Sequence=Displayed;
CC Name=Sema Y-S;
CC IsoId=O9WLT3-2; Sequence=NSP_066048;
CC 4. TISSUE SPECIFICITY: EXPRESSED IN MANY REGIONS OF THE DEVELOPING
CC NERVOUS SYSTEM, PROBABLY IN NEURONS AND THEIR PRECURSORS, BUT ALSO
CC IN NONNEURAL TISSUE SUCH AS IMMATURE MUSCLE AND DERMIS. IN ADULT,
CC STRONG EXPRESSION IN THE SKELETAL MUSCLE AND MODERATE EXPRESSION
CC IN THE BRAIN, WHERE CEREBELLUM SHOWS THE HIGHEST EXPRESSION. ALSO
CC EXPRESSED IN ALMOST ALL AREAS OF THE CNS.
CC 5. DEVELOPMENTAL STAGE: DETECTED AT E1.2 AND FOUND AT MARKEDLY
CC INCREASED LEVELS AT E1.5 AND E18 IN BOTH THE HEAD AND THE BODY. AT
CC BIRTH THE LEVEL DECREASES SIGNIFICANTLY.
CC 6. SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC 7. SIMILARITY: Contains 1 Sema domain.
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB003817; BAA76293.2;
CC EMBL: AB014074; BAA76295.1;
CC InterPro: IPR001627; Sema.
CC Pfam: PF01403; Sema; 1.
CC SMART: SM00630; Sema; 1.
CC Signal: Transmembrane; Multizone family; Neurogenesis; Glycoprotein;
CC Developmental protein; Alternative splicing.

```

```

I  SIGNAL  1  23  POTENTIAL
I  CHAIN  24  960  SEMAPHORIN 60 (POTENTIAL)
I  DOMAIN  24  635  EXTRACELLULAR (POTENTIAL)
I  TRANSMEM  636  636  POTENTIAL
I  DOMAIN  657  960  CYTOPLASMIC (POTENTIAL)
I  DOMAIN  234  541  SEMA
I  DOMAIN  693  639  POLY-PRO
I  DOMAIN  783  736  POLY-PRO
I  CARBOHYD  69  459  N-LINKED (GLCNAC) (POTENTIAL)
I  CARBOHYD  285  285  N-LINKED (GLCNAC) (POTENTIAL)
I  CARBOHYD  436  436  N-LINKED (GLCNAC) (POTENTIAL)
I  VARSPLIC  586  617  Missing in isoform Sema V-S1
I  /FTID=VSP 006C48
I  SEQUENCE 360 AA: 1026:0 YW; C88233C560726086 CPC44;
Query Match 27.4%; Score 1491; DB 1; Length 960;
Best Local Similarity 35.1%; Pred. No. 31e-88;
Matches 374; Conservative 149; Mismatches 363; Indels 180; Gaps 32;
I  6 LLLYFLLHFAGAGPEDESEPSSSHGNTYKQYVYVGHKPGENTTCRHRLDIQNMINK 65
I  11 LLLLLLSCQAQTAFPCDIPLETSDQSTSSSFRGJEDDAVAHELGLSFOBLTJN 69
I  66 GLLYIAARDHIYVDIDTSTEEIYCSKLTWKSQAQVETCRVKGFKDCRHNFIKVL 124
I  70 RLLYVAARDHVFDFQAGEEGELVPKXELTWSC--DNVCAVRGKLTDECVNYIVL 127
I  125 LKNDALFVCGTNAFNPSCRYMYMDTBEFQDESGMARPCYDAKHAVALFADGKYS 184
I  128 VFWDSQTLACGTTNSFPCRSYCYTSLDCEELSGQARCPDTQSTVAISAEGSYVS 187
I  185 ATVTDFLALDAVYRLSLGSPSLFTYKHSKVLKBPYFVCAVDVGGYVYFFREYAV 244
I  188 ATAADPQASDAVYRLGQPLFSAKYSKWLRRHRFVALHGHCHVYFFREYAVSICA 247
I  245 TMGKVYFPPVAQVKNDMGSSRVLEKQWTSFKARLCKSVSGSGHVENILOAVTYR 304
I  248 RLGRVQFSRAVAVCKDMGSSPRALDEHTSPFKRLNGSVFSGSTFFEDVQLSLGPV 307
I  305 IIGRDVVLATFTPYNSIIGSAYCAVNDKIDIASVFTGRPKSKSGSTWTFPDERVYK 364
I  308 LHGRSALFGVFTTQNSIPSSAVCAFEVLECIPEGPBKGKESGLDGMATFVEDKVPSP 367
I  365 RPCCAGSSSLERVATSNFPEDDLAFIKTRPMDAVESIFENKWFRTWRYELTIA 424
I  368 RQSCAGVGAALFSSQQLPDLFLFKARPLRPAPVPAHOPMLTLRALLTQVA 426
I  425 VDTAGPYQNTVYVFGSEKGIKFLKARIGNSGFLNLSLFEKSVYNSKSYDGVED 484
I  427 VQMGAGFHRNTTVFLGSDGTVLKVPSP--GGSLGPERITLDEIDAYSHARCS--GKRS 483
I  485 ----KEINGVCLDRASSLYVAFTCVIYVPLGRCEHGKCKKCTIASDPYCGMKLGG 540
I  484 PRAARIIGLELDEGHRUFVAFEGGIVYLSRCAHAGACORSCLASDPYCGHRRFG 543
I  541 ACSHLSPKSLRTEQDIE RGTMG--GDCHNSFVALN---GHSSSL-----PSITT 586
I  544 CYNIRPFG-----TVDLTGNCSENEHCCCGGATGSGSGGSDAYVLLGPGPPEPS 598
I  589 SDSTAQEGYEGSGMLDWKHLDSPOSTPLCAVSSHHNQKGVIRESVYKXGHPQVVP 649
I  599 SPSCAHFGPGS-----STLGA-----HCGVPRDLSPASASGSP 634
I  649 TLIAIAVLAIFUNGAVFSGITVYVCEHRAKQVAVVORKEKLTSHRRGSMSTVK; SGL 708
I  635 PLLIACVAALFALGASVSGLVNSCAG--RAN-----RRSKDIEPLGPRP; SLRSLAR 688
I  709 FGTQSKDKPP--EAILPLXHGKGLATPGNTAKVLKADQHHLDTALPTPESTPILQ 763
I  689 HGGGPPEPPPKGDEAAQTPQLVITTLPPPEGSGPP-----ELACUPTPTTPELP 739
I  766 QKRKPSRGSEWRPNONINACTKDMPPKSGSVIPTCLP---LRASPSHIFSVVYVLPITQ 822

```

## RESULT 5

SMIA\_SCHAY

```

ID SMIA_SCHAY STANDARD: PRT: 730 AA.
AC Q26473;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 1A precursor (Semaphorin-I) (Sema I) (Fasciclin IV).
GN SEMA-1A CR PAS4.
OS Schistocerca americana (American grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
CX NCBI_TaxID=7009;
RN 1;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91C40225; PubMed=1418998;
RA Kolodkin A.L., Matthews D.J., C'Connor T.P., Patel N.H., Admon A.,
RA Bentley D., Goodman C.S.;
RA "Fasciclin IV: sequence, expression, and function during growth cone
RA guidance in the grasshopper embryo.";
RL Neuron 9:83-845(1992).
CC -! FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
CC -! SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -! TISSUE SPECIFICITY: DYNAMICALLY EXPRESSED ON A SUBSET OF AXON
CC PATHWAYS IN THE DEVELOPING CNS AND ON CIRCUMFERENTIAL BANDS OF
CC EPITHELIAL CELLS IN DEVELOPING LIMB BUDS.
CC -! SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -! SIMILARITY: Contains 1 Sema domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EXBL: L00709; AAA29808.1;
CC PIR: JH0798; JH0798.
CC DR InterPro: IPR03659; plexin-like.
CC DR InterPro: IPR002165; plexin_repeat.
CC DR InterPro: IPR01629; Sema.
CC Pfam: PFG1437; PSI;
CC Pfam: PFG1403; Sema;
CC Pfam: PFG1403; Sema;
CC SMART: SM00423; PSI;
CC SMART: SM00633; Sema;
CC Signal: Developmental protein, Transmembrane; Glycoprotein;
CC Neurogenesis.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 730 SEMAPHORIN 1A.
CC DOMAIN 21 630 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 631 651 POTENTIAL.
CC DOMAIN 652 730 CYTOPLASMIC (POTENTIAL).

```



```

T DOMAIN 215 515 SEXA.
T DOMAIN 217 222 POLY-PHE.
T CARBOHYD 44 44 N-LINKED (GLNAC...) (POTENTIAL)
T CARBOHYD 71 71 N-LINKED (GLNAC...) (POTENTIAL)
T CARBOHYD 163 163 N-LINKED (GLNAC...) (POTENTIAL)
T CARBOHYD 267 267 N-LINKED (GLNAC...) (POTENTIAL)
T CARBOHYD 360 360 N-LINKED (GLNAC...) (POTENTIAL)
T CARBOHYD 539 539 N-LINKED (GLNAC...) (POTENTIAL)
O SEQUENCE 730 AA: 8,214 MW: 626259467E7F8E57F CRC64:
Query Match 17.83; Score 971.5; DE 1; Length 730;
Best Local Similarity 34.43; Pval: No. 5.9e-55;
Matches 244; Conservative 124; Mismatches 224; Indels 127; Gaps 22;
Y 8 LYFTLLHFAAGPPEDPEP-SISGNNTKYPYFVGHKRCRNTQP-----RLDIOVI 61
b 11 LLWALH--AAAWNDUSP-----KVVQF-----GEERVQDFLNFSGKCHFKL 54
Y 62 MIMGTUYLAARDHIYVD-D--TSHTETTYGSKKLTWSPQADVNTQYKHGKHDECHN 119
b 55 EKDKNSLLVGARNVWNSLSRDLTEPTQ-----RTEWSSGAEARELCILKQSESDYCN 109
Y 120 FIKVLKKQDAIPVGCNAPKPSGNYVDTLEPQD ---EPGVARCPYCAKHWY 174
b 110 YRVLAIEDRVLCOTNAYKPLCRHYALD-----GVYVWEYETGRGLCFDPDHNHT 169
Y 175 ALFAEGRLYSATVDFATCAVIVSLSGESLTLTQYHNSKMKKEFYFQADYDNYTFP 234
b 166 AIYEGGLYSATVADFGSTDLIVFG-----PLTERSDAKQLNAFNVTWYEDCFIFP 220
Y 235 FFRSLAVYNTGKVPFPAVACVKNMGSCQVLEKCHTSFKALNVSQVDSHFVEN 294
b 221 FFRATAYEYVNOGKAIVARVCKDKGHQFQGRP-WTSPILKSLNCSVGGDYFFYFN 279
Y 295 ILOAVTDVIREN-GRDY-----VLAIFSPFYNKISGSAVCAVCMULIASVTGFRKESKD 350
b 280 EICSTSLIEGNGGVQVEKLIYGVFTTPNKSIGSAVCAESMSKLSLEDGFRKECTYN 339
Y 351 STWTPVDPERVPRPCGAGSSLEKATSNFPDPLNFITKPLHMTAEPSPFKRW 410
b 340 SNWLAVPSLKVPRPQGVND-----SRLPLDVSNVFNKSHYTLDEAVFAFTRPI 391
Y 411 FLRTMVRPLTKIAND-----TAGPYQNTVNLGEGKGLILAFIARIGNSGFLN----- 461
b 392 LIRLSGRFKIADQCQVTPQD--KAYDVLFGIDLGKVIKAL-----NSAFSDSDTV 445
Y 462 DSLFLEKSYNNEKSGYGVDEKRYGMOLDRASSSLVYAFSTCVIKVPLQRC--EREG 519
b 446 DSVVIEELQVLP-----PGVPVNLVWMDGSDSKLVVSDDELAIKLHRCGSDKIT 499
Y 520 KCKTKTASRPYCGNLIKESGAGHL-SPN-----SKUTFEQDIERGNTDGLGDCHNSFVAL 575
b 500 NCRE-CVSLQDPYCAWNLKCTAVGSPQWAKGRFFIQNLSLGEKACGGPQCEIV- 557
Y 576 NGRSSSLPSTTSQSTA-----QGVESRGGXLDKWKHLSDSPSTPLGAVSSRHHQ 626
b 558 ----ASPVPPTTKSSCPVSHIQHAEFPE---IDNEIVIGVDSNVIPRTLALINHA 610
Y 629 DKGVIRESYLKHGDDQVVP-----TL-LAI-----AVLIAFYGVGAVFS 666
b 611 GSK-----LPSQEKLPITYAETLTATVTSIGLGAJVVGF-SGRIFPS 652
O SEQUENCE 730 AA: 8,214 MW: 626259467E7F8E57F CRC64:

```

ESTUT 10

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

DIA DROKE

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

[1]

SEQUENCE FROM N.A.

TISSUE=Embryo;

MEDLINE=94094332; PubMed=6269517;

Kolchkin A.L., Matthews D.J., Goodman C.S.;

"The semaphorin genes encode a family of transmembrane and secreted growth cone guidance molecules";

Cell 75:1389-1399(1993);

SEQUENCE FROM N.A.

MEDLINE=20196006; PubMed=1071132;

Adams K.B., Celiker S.E., Holt R.A., Evars C.A., Scayne C.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman G.R., Vardell M.C., Zhang Q., Chen L.X., Branson A.C., Rogers Y.-H.C., Blazer P.G., Champe M., Pfeiffer B.D., Wan X.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., McKus G.L.G., Abell C.F., Agbayani A., An H.-C., Andrews-Pfannkoch C., Baldwin D., Ballew A.Y., Basu A., Bayazitov I., Bayraktarov G., Beasley E.M., Beeson K.V., Benos P.V., Bertan S.P., Bhandari D., Bolshakov S., Bonkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P., Burris K.C., Busch D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de la Fabia J., Delecher A., Deng Z., Mays A.D., Bew I., Dietz S.X., Dodson J., Doup E.E., Downes M., Dugan-Rocha S., Durkin B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foeller C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K., Glendon A., Gong F., Gorrell C.H., Gu Z., Guan P., Harris M., Harris K.L., Harvey D., Heiman T.C., Hernandez C.R., Hock C., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwari C., Jaitani Y., Kalush F., Karpen G.H., Ke Z., Kennison C.A., Ketchum K.A., Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko F., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Maitte B., McIntosh T.C., McJeed M.P., McPherson B., Merkulov G., Milshina N.V., Nobury C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M., Palazzolo J., Pittman G.S., Pan S., Pellaard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shie B.C., Siden-Kiamos S., Simpson M., Skupski X.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

The genome sequence of *Drosophila melanogaster*;

Science 287:2185-2195(2000).

FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.

SUBCELLULAR LOCATION: Type : membrane protein.

TISSUE SPECIFICITY: EXPRESSED BY SUBSETS OF NEURONS AND MUSCLES.

DEVELOPMENTAL STAGE: EXPRESSION BEGINS AROUND STAGE 13, PRIMARILY IN THE DEVELOPING CNS. IN STAGE 16 EMBRYOS, IT IS EXPRESSED AT HIGHER LEVELS THROUGHOUT THE CNS, AND WEAK EXPRESSION IS SEEN IN PORTIONS OF THE PERIPHERAL NERVOUS SYSTEM, MOST CLEARLY IN THE LATERAL SENSORY CLUSTERS.

SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

SIMILARITY: Contains 1 sema domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.sib-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL: L26082; AAA88789.1;

```

EMBL: AF003621; AAF52695.1; ..
FlyBase: FBgn0011259; Sema-1a.
InterPro: IPR003659; Plexin-like.
InterPro: IPR002165; Plexin_repeat.
Pfam: PF01437; PSI; ..
Pfam: PF01403; Sema; ..
SMART: SM00423; PSI; ..
SMART: SM00423; Sema; ..
Signal: Developmental protein; Transmembrane; Glycoprotein;
Neurogenesis.
SIGNAL 1 19 POTENTIAL
CHAIN 20 771 SEMAPHORIN 1A.
DOMAIN 21 606 EXTRACELLULAR (POTENTIAL).
TRANSMEM 605 629 POTENTIAL.
DOMAIN 630 771 CYTOPLASMIC (POTENTIAL).
DOMAIN 630 771 SEMA.
DOMAIN 689 694 POLY-PRO.
CARBOHYD 41 41 N-LINKED (GLCNAC. ) (POTENTIAL).
CARBOHYD 68 68 N-LINKED (GLCNAC. ) (POTENTIAL).
CARBOHYD 138 138 N-LINKED (GLCNAC. ) (POTENTIAL).
CARBOHYD 158 158 N-LINKED (GLCNAC. ) (POTENTIAL).
CARBOHYD 262 262 N-LINKED (GLCNAC. ) (POTENTIAL).
CARBOHYD 355 355 N-LINKED (GLCNAC. ) (POTENTIAL).
SEQUENCE 771 AA; 86935 MW; 75267005.6F330DAS CRC64;
Query Match 17.1%; Score 932.5; DB 1; Length 771;
Best local similarity 35.0%; Pred. No. 2.1e-52;
Matches 219; Conservative 136; Mismatches 233; Indels 87; Gaps 20;
57 TLVIAARHGYVTDITSTHTSEIVSGKLTWKSADVDTCRKGKKGKDECHNFIKVLK 126
57 SLLIGARNTVFHLS-----HCLVQQRLLWTSPEDDTKACLVKKGSEACONVIRVW 111
127 KNDALFVCGTNAFNSCRNYKHC-----TLSPFGDPSPGSMARCPYDAKHANVLPADSKL 182
112 PEPGRFLVGTGNSFRPMCHNTYIISDSNYTLEA-----TKGQAVCPYDPHNSVSLADNEL 168
183 YSATVTDFALDAVIYRSGESPTLVTHDKDKLKERFYQAVDYGIVYIPFPRRIAVE 242
169 YSGTVADFGSDPFIYRE-----FQTEQYDLSLNATNFVSEFTCGDPVFFPFFETAVE 223
243 YNTMGKVFPRVAQCKNDKNGGSCQVLEKCMTSFLKARNTGVPDGHSHFYFLICAVTCV 302
224 FVNCSEKAVSRVARCKWMDKGGFER-FRNRWTSFLKSRNLNCSIPGDPFFYNEISASNL 282
303 IR-----INGRDYVATFTPYNRSFGSAVCAYDNLEIASVFTGRFKGKSPSTWTPYED 358
283 VEGQYGMSSKLIYGVNTPSKSIPEGAVCAPALGCIATTFEQGKEQTGNSNMLPKNN 342
359 ERYKPRPCGCGAGSSS:RYATSNRPPTDITNPTKTHLMRDAVSSINRPFVFLRMTRY 418
343 AKVPDPFGSC-----HNFSFALPDFTLNFKTHSLXNDENVFAPFSQPIAKRTSTV 394
419 RTTKIAVD---TAAGEVGNHTVAFLGSEKGIILKFLAFIGNSGFINDSLPLEBMSVYN 473
395 RTQIAVDAQIKTPG--KTYDVIVGTDHGLKIISYNNESAGSKVTSVVISIDVLT 452
474 SEKCSYGVGVEDKRI:GMQLDRASSLSY-----VAFSTCVIKVPLGRC--ERHSGKOKRT 524
453 KS-----EPIRNLIVATMQVDQPDSSYDQGLIIVTDSQVVAIQHRCNDKITSQSA- 507
525 CIASRDFYCGWKEGAC-SHLSFN---SLTTEQDIERGNTDGLGCHNSFVALNGHSS 561
508 CVALQDFYCAWOKIACKRSHGAPRLEENFYQNVATCG-----HAAC 551
582 LLPSTTTSSTAQGVESGGLWKKHLLDSCDSTPLGAVSSHSHQDKGVARESYLKG 641
552 PGKINKSKANAGEQKGFNDM-----CILDS-----RRQSKQOEIION:DKN 594
642 HQQLVPVTLIAIVILAFNVGAVFS 666
595 FDIINAQYVTEILWAVLAGSIFS 613

```

```

RESULT 11
SMA_RAT ID SMA_RAT STANDARD: PRT: 772 AA.
AC Q63548;
DT 30-MAY-2000 (Rel. 39, Created);
DT 30-MAY-2000 (Rel. 39, Last sequence update);
DT 15-SEP-2003 (Rel. 42, Last annotation update);
DE Semaphorin 3A precursor (semaphorin III) (Sema III).
GN SEMA3A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN 1.
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=9707089; PubMed=8915827;
RA Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaagen J.;
RT "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
PT relationship to developing nerve tracts during neuroembryogenesis.";
RL J. Comp. Neurol. 375:378-392(1996).
CC 1- FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
CC NEUROFILIN.
CC 1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC 1- DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELECEPHALIC
CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
CC SOVITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
CC MITAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
CC CEREBELLAR PURKINE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
CC SIGNAL MOTONEURONS.
CC 1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC 1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC 1- SIMILARITY: Contains 1 Sema domain.
CC 1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch.
CC
CC EMBL: X95286; CAA64607.1; ..
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003026; Ig_XHC.
CC InterPro: IPR003659; Plexin-like.
CC InterPro: IPR001627; Sema.
CC Pfam: PF00047; Ig; 1.
CC Pfam: PF01403; Sema; 1.
CC SMART: SM00429; IG; 1.
CC SMART: SM00423; PSI; 1.
CC SMART: SM00423; Sema; 1.
CC SMART: SM00630; Sema; 1.
CC PROSITE: PS00835; IG LIKE; 1.
CC Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
CC Developmental protein; Glycoprotein.
CC SIGNAL 1 220 POTENTIAL.
CC CHAIN 21 772 SEMAPHORIN 3A.
CC DOMAIN 240 538 SEMA.
CC DOMAIN 577 665 IG-LIKE C2-TYPE.
CC DOMAIN 728 770 ARG/LYS-RICH (BASIC).
CC DISULFID 650 723 BY SIMILARITY.

```

```

T CARBOHYD 53 53 N-LINKED (GLCNAC... (POTENTIAL)
T CARBOHYD 125 125 N-LINKED (GLCNAC... (POTENTIAL)
T CARBOHYD 591 591 N-LINKED (GLCNAC... (POTENTIAL)
Q SEQUENCE 772 AA: 88808 MW: 240907612FF9F2D2 CRC64:
Query Match 15.78; Score 856; DB 1; Length 772;
Best Local Similarity 32.38; Pres. No. 1.3e-47;
Matches 209; Conservative 104; Mismatches 25; Indels 54; Gaps 20;
47 GRNTQRRHLDICVIMINRG...LVIAARDHVTVDICT 83
25 GKNNVPLKLSYKEMLSNNVITENG...ANSSVHTFLDEERSPLVGCARDHFSN... 82
84 SHREELVCSKRLTKWKSQADVCTRMGKH...KECHNFIVYLLKJKNDDALFVCGTNAFNP 142
83 ---VNKDFGKIVFVSVYTRDECKWAGKILKPCANTIKVLKAYHOTHLYAGCTGAERP 139
143 SCR-----NVKVCOTLEPTGDEF--SGWARGPYAKAKVALFADGKLYSATVDFLADA 195
140 ICYVIEVCHFEENIFLQSHENGKGPYDFKJLTASLLDGENVSTADPFKQDP 195
196 VYRSGESTLTIVKHSKWLKEPYVQA-----VQYGDYIVFFFEZAVEVNTMGX 248
200 AIPETLGHHPIRTEQDSKRLNDPRISAHILIPESNDPEDEKVVFFFRNADGHSOK 253
249 VVFPVACVKNMGGGRVLEKQWTSFLKARLNCVFG--DEHFYFNLQAVTVYIN 304
260 ATVARIGCIKNDFFGG-HRSVYKWTFLKARLICSVEFGNGLDTHF-----DELDQVFL 418
305 INGRD---VYLAFTSTPVKSIHSAVCAVEMLIASVFTGREKCKSPDSTVTPQSR 360
314 MNSKQPPVYGVFTSSNIFXSAYVMYKQDVRFVGLPVHREGNYQWVF-QGR 392
361 VPKRPCCAGSGSLRYASNBEFTDLTFKTHIMDSAVPSIFNRWFLFVYVR 420
373 VVYPRPQTCR-SKTEGDFDSTKQLEDDVITPASHSPFMVPEVPIKNERLYKNTDNYOF 451
421 TKIADVTAGPYQNTVTVVLGSEGIILKLAHLKSGFNDLSPLENSVYNSKCSYR 490
432 TQIWDVDAEDQGVYFVIGTVGTVLKVVSVKPEKWHDELVVLEEMTVFR----- 494
491 GVEDKRIYNGQLDRASSLLVAFSTCVIKYPLGRCERHGHCKKCTCIASRDEYCGWKEGG 540
465 --EPTTISAMELSTKQQLIGSTAGVAGQPLHRCQYGRACRECLARDFYCAW--DGS 540
541 ACHLSLSPNS-RLTFEODIERGNTQGLGCC-----HNSFVALNGHS--SSLPSTTTSDST 592
541 SCRYFPPTAKFRTRRQDTRNG--DPLTHCSDLQHNCHN---HGHSLERKLYGVENSSTF 595
593 AQEGYSGRGYLDWKH-LDSPDSFDPLGAVSSHHCQKQKQVIESVLK 640
596 LECSPKQALVYQVQFRRNEDRKEEI-RVGDH:::RTEQGLLRSIQCK 642

```

## RESULT 12

```

M3A_MOUSE
D SW3A_MOUSE STANDARD; PRT; 772 AA.
C Q08665; 362180; C62215;
T 30-MAY-2003 (Rel. 35, Created)
T 30-MAY-2003 (Rel. 39, Last sequence update)
T 15-SEP-2003 (Rel. 42, Last annotation update)
E Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D'
E (Sema D)
N SEMA3A OR SEMA3 OR SEMD.
S Mus musculus (Mouse).
S Eukaryota; Metazoa; Chordata; Crniata; Vertebrata; Euteleostomi;
X Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.
X NCBI_TaxID=10090;
N [1]
P SEQUENCE FROM N.A.
S STRAIN=MNRI; TISSUE=Embryo;
X MEDLINE=95267431; PubMed=7748561;
A Pueschel A.W., Adams R.H., Betz H.;

```

```

RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RT and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97473885; PubMed=9331345;
RA Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
RA Yagi T.;
RT "Disruption of semaphorin III/D gene causes severe abnormality in
RT peripheral nerve projection.";
RL Neuron 19:519-530(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX Kimura T., Fishman M.C.;
RA "cDNA sequence of mouse collapsin/semaphorin III.";
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 107-772 FROM N.A.
RX "SSEA fetal brain;
RX MEDLINE=95267432; PubMed=7749562;
RA VetterSmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
RA Goodman C.S., Kolodkin A.L.;
RT "Semaphorin III can function as a selective chemorepellent to pattern
RT sensory projections in the spinal cord.";
RL Neuron 14:649-659(1995).
CC -!- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
CC PATTERN SENSORY PROJECTIONS BY SELECTIVELY REPELLING AXONS THAT
CC NORMALLY TERMINATE DORSALLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
CC -!- IN DISTINCT REGIONS OF THE NEUR ectoderm AND MESODERM.
CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X85993; CAA59385.1;
CC EMBL: D85028; BAA1973.1;
CC EMBL: A41541; AA177612.1;
CC EMBL: J40484; AAA73934.1;
CC PIR: I48747; I48747.
CC PIR: I58163; I58169.
CC MGD: MGI:137558; Sema3a.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003599; IG.
CC InterPro: IPR03006; IG_MHC.
CC InterPro: IPR03659; Plexin-like.
CC Pfam: PF01403; Sema; 1.
CC SMART: SM00409; IG; 1.
CC SMART: SM00423; PSI; 1.
CC SMART: SM00630; Sema; 1.
CC PROSITE: PS50835; IG LIKE; 1.
CC Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
CC Developmental Protein; Glycoprotein.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 772 SEMAPHORIN 3A.
CC DOMAIN 240 539 SEMA.
CC DOMAIN 579 665 -IG-LIKE C2-TYPE.
CC DOMAIN 728 773 ARG/LYS-RICH (BASIC).
CC DISUFID 650 723 BY SIMILARITY.
CC CARBOHYD 53 53 N-LINKED (GLCNAC... (POTENTIAL).

```

```

T CARBOHYD 125 N-LINKED (GLCNAC... (POTENTIAL)
T CARBOHYD 191 N-LINKED (GLCNAC... (POTENTIAL)
T CONFLICT 193 D -> N (IN REF. 4)
T CONFLICT 207 H -> D (IN REF. 1)
T CONFLICT 253 D -> G (IN REF. 1)
T CONFLICT 352 F -> G (IN REF. 4)
T CONFLICT 403 A -> G (IN REF. 1)
T CONFLICT 571 OH -> ED (IN REF. 1)
T CONFLICT 616 EDRKE -> RSKR (IN REF. 1)
T CONFLICT 623 P -> K (IN REF. 4)
T SEQUENCE 772 AA; 88799 MW; E69A0852B5-0AFC3 CRC64;

Query Match 15.7%; Score 253; DB 1; Length 772;
Best Local Similarity 32.1%; Pred. No. 2,9e-47;
Matches 208; Conservative 102; Mismatches 257; Indels 78; Gaps 18;

Y 47 GRNTTQRRLDQIMINMGNT-----LYIAARDHIYTVDT 83
| | | | | | | | | | | | | | | | | | | | | |
D 25 GKGVPRLKLSYKEMLESKNVITFNGLANSSVHTLLDERSRLVVGAKHIFSFNL-- 82
| | | | | | | | | | | | | | | | | | | | | |
Y 84 SHTEEYCSKKLTKNSROADYDTCRKGKH-KDECHNFKIVLLKNDDALFYCGTNAFNP 142
| | | | | | | | | | | | | | | | | | | | | |
b 83 ---VNIKDFQKIVWPVSYTRRDECKWAGSDILKECANFKIVLEAVNQTHLYACGTGAPRP 139
| | | | | | | | | | | | | | | | | | | | | |
Y 143 SCR-----NKKXDTLEPFGEF--SCMACPYDAKHANVALFADGKLYSATVTDPLAIDA 195
| | | | | | | | | | | | | | | | | | | | | |
D 143 ICTYISVGHHPEDNIFKLQSHFNGRSGSPDPKLLTASLLIDGELYSGTANDWGRDP 199
| | | | | | | | | | | | | | | | | | | | | |
Y 196 VYRSLSGESPTLTAVHDSKWLKEPYEVA-----VQYGYIYFFFEERAEVNTMOK 249
| | | | | | | | | | | | | | | | | | | | | |
D 200 AIFRTLGHHPITRTEQHSKWLNDPFIHAHLIPESCDNPDCKWVFFPENAIDGHSKK 259
| | | | | | | | | | | | | | | | | | | | | |
Y 249 VVPRVAQCKMDGMSORVLEKQWTFKARLNCVFG-----DSHFYENILCAVTDVIR 304
| | | | | | | | | | | | | | | | | | | | | |
D 260 ACHARIGQCKNDPFG-HRSIYKWTFTFKARLICSVPGRSGDTHF-----DEGCVPE 313
| | | | | | | | | | | | | | | | | | | | | |
Y 305 INGRQ-----VVIATFSTPYNSIPGSAVAYCMLDTASVFTGPKCKSPDSTWTPVPER 360
| | | | | | | | | | | | | | | | | | | | | |
b 314 MNSKDPKNP-VYGVFTSSNIFKSAVMSVKSQDVRVELGFAHEDGNYQWVPY-QGR 372
| | | | | | | | | | | | | | | | | | | | | |
Y 361 VPXPRGCGAGSSSLERYATSNFPDCTINFETHPLMDAEVPSIPNRWFTRTWVRYEL 420
| | | | | | | | | | | | | | | | | | | | | |
b 373 VPYPRGTCP-SKTFGFGDSTKLPDEVITEARSHPAVNPVFFINRIM-KTDVYQCF 431
| | | | | | | | | | | | | | | | | | | | | |
Y 421 TKAVDTAGAPYGNHNVVLGSKGILLKPLRIGNSGFLNLSPLFEESVYNSEKSCVD 480
| | | | | | | | | | | | | | | | | | | | | |
D 432 TQVDRVDAEDCCYDMF-GTDVGVLKVVSVKETHWCHLEVILSEVTVPR----- 484
| | | | | | | | | | | | | | | | | | | | | |
Y 481 GVEDKRMGMQLDRASSSYVAFTSTVYKVPACRCHREKCKKTCIASRDPYCGNKEGG 540
| | | | | | | | | | | | | | | | | | | | | |
D 485 --EPTTISAMELSTRCCCLYIGSTAGVAILP-HRCDIYKKAACCLARCPYCAN--DGS 540
| | | | | | | | | | | | | | | | | | | | | |
Y 541 ACSHUSPNS-PUTTEQIERGNTGGLGCHNEFVALNGHSSS-----JSTSTTSSTAGE 595
| | | | | | | | | | | | | | | | | | | | | |
D 541 SCRSYPTAKRRTRRCQIRNG--DPLTHCSOLQHDHWHGHSLEERIIYGVENSSTFIFC 598
| | | | | | | | | | | | | | | | | | | | | |
Y 596 GYSGRGMLDKWILLDSPSTDFLGAVSSNHQKGVTPRESIK 640
| | | | | | | | | | | | | | | | | | | | | |
D 599 SPKSQALVYVQFQRNEQRKEE-RYGDHIITFEQGLLINS-LQK 642
| | | | | | | | | | | | | | | | | | | | | |

RESULT 13
M3A_HUMAN
D_M3A_HUMAN STANDARD; PRT; 771 AA.
C Q14563;
T 30-MAY-2000 (Rel. 39, Created:
T 30-MAY-2000 (Rel. 39, Last sequence update:
T 15-SEP-2003 (Rel. 42, Last annotation update:
E Semaphorin 3A precursor (Semaphorin I.1) (Sema I.1).
S SEMA3A.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```

NCBI TaxID=9606;
(1)
SEQUENCE FROM N.A.
TISJB=fetal brain;
MEDLINE=94094332; PubMed=5269517;
RA Kolodkin A.L., Matthes D.C., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
growth cone guidance molecules.";
RL Cell 75:1389-1399(1993);
(2)
SEQUENCE OF 1-37 FROM N.A.
RA Moessner J., Vintz P., Hinds K., Strommatt C.;
PL Submitted (MAR-1998) to the EMBL/GenBank/DBS databases.
(3)
SEQUENCE OF 39-182 FROM N.A.
RA Rosling T., Tih-Woliam A.M., Duckels G.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBS databases.
CC !- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
NEUROFILIN-1/PLEXIN-1 (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: Secreted (by similarity).
CC !- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
THIRD OF THE PROTEIN.
CC !- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC !- SIMILARITY: Contains 1 Sema domain.
CC !- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/a/announce/
or send an email to license@isb-sib.ch).
CC
EMBL: L26381; AAA65338.1;
EMBL: AC004451; ; NOT ANNOTATED_CDS.
EMBL: AC004848; AAC78622.1;
PIR: D49423; D49423.
Genbank; HGNC:10723; SEMA3A.
M3V; 603961;
GO; GO:0005576; Cytoextracellular; TAS.
InterPro; IPR007110; Ig-like.
InterPro; IPR003599; Ig-like.
InterPro; IPR003006; Ig-MHC.
InterPro; IPR003659; Plexin-like.
InterPro; IPR001627; Sema.
Pfam; PF00047; Ig_1.
Pfam; PF01403; Sema_1.
SMART; SM00409; IG_1.
SMART; SM00423; PS_1.
SMART; SM00630; Sema;
PROSITE; PS00835; IG-LIKE;
Signal; Immunoglobulin domain; Multigene family; Neurogenes.3;
Developmental protein; Glycoprotein.
FT SIGNAL 1..20 POTENTIAL.
FT CHAIN 21..771 SEMAPHORIN 3A.
FT COMAIN 240..538 SEMA.
FT COMAIN 580..664 -IG-LIKE C2-TYPE.
FT COMAIN 727..769 ARG/DYS-RICH (BASIC).
FT DISULFID 649..722 BY SIMILARITY.
FT CARBCHYD 53..53 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBCHYD 125..125 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBCHYD 590..590 N-LINKED (GLCNAC... (POTENTIAL).
SQ SEQUENCE 771 AA; 88889 MW; 9985F8D3EAD8456 CRC64;

Query Match 15.6%; Score 852.5; DB 1; Length 771;
Best Local Similarity 32.1%; Pred. No. 2,9e-47;
Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;

QY 44 HKPGRNTQRRLDQIMINMGNT-----LYIAARDHIYTVDT 80
| | | | | | | | | | | | | | | | | | | | | |
```

```

b 22 YQCKXNVPRKLSYKMLSENNVITFNLGLANSSVATFLDZERSELYUGAKDHIESFD 81
b 81 IDTSHTTEIYCSKLLWKSKQACVDTCKMKGR-KQCHNFIFKVLKQDDA-FVCGTNA 129
b 82 L-----VNIKDPQKLVMPVSVTRDECKWAGKQILKSCANFIRKVKAYNCTHLYACGTGA 136
b 140 ENRSCR-----NKKMUTLEFGDF--SGMARCEYAKAHANVALFADGKLYSATVTDFLA 192
b 137 FHCITYEIGHPEPDIKLENSHFENGKSPYDPKLTASLLDGLGYSTAADEMG 196
b 193 IDAVIRSLGESPTLRVVKDSKALKPVPVQA-----VOYGVYVFFFRDIAEYNT 245
b 197 RDAIFPTLGHHPHPIRTECHDSRLNDPKLSAHLSESNDPDKVYFFFRNADGEH 256
b 246 MGKVVPFPAQVQNVKGGSCRVLEKQWTFELKARLNCVFG-----DSHFYFNLQAVTD 302
b 257 SGKATHARIQICKNDFGG-HRSJVNKWTFLKARLICSVEGPKG-DTHF-----DELQD 310
b 302 VIR-NGRD---VVIATFSPYKSIKPSGAVCAVDYLDIASVFTGRFKEQKSPDSTNPF 357
b 311 VFLKNFQPNFVYGVFTTSSNIFKGSVAVCYMSDVRVFLQPAHRDGPYQWVPY 369
b 359 DERVPKPRPCCAGSSSRRYATSNFPPDITPIKTHPLVDRAVESIFKPFMFRTMYR 417
b 370 QGRVYFPGTGP-SKTFGGFDSKCLPDGVITFAPSHIPAYVNPFPMPNRPVINTDM 428
b 418 YRLTKIANDTAGPYQNHVYVLSSEKGIILKLARLGNSEGLNDSLFLEKSNYNSKC 477
b 429 YQFTQVYVDAEDGGYDNYFISTDGVTVKVVSTFKETVYGLRVLSEMYVFP-----484
b 478 SYDGVEDKRMGMCLDRASSLYVAFSTCVIKVPLGRCEHGRKCKTCTIASRDPYQGNIK 537
b 485 -----EFTASAMELSTKQQLYIGSTAGVAQPLHRCDIYGVKACAECLAPRYCAM-- 537
b 538 EGGACSHLSPNS-ELTTEQDIERANIDELQD-----HNSFWALNRS--SSLLPSTTSD 592
b 538 DGSACSRYPFTAKRTRRQDINQ--DELTHCSLHEDN--HGRSFEKFIYGVENSS 592
b 591 STAGEVYESGGMLDAKHLDLSPSTPFLGAVSHNHQKGVIFER 637
b 593 FTLECPKSGACVYQWQPRNEERKEEI-RVDEHIIRTDGGLLRS 638

```

## RESULT 14

```

SMIA-TRICP STANDARD: PRT: 712 AA.
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2000 (Rel. 40, Last annotation update)
Semaphorin-1A precursor (Semaphorin-1A)
SEMA-1A OR SEMA-1
Tribolium confusum (Confused flour beetle),
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
NCBI_TaxID=7071;
11.
SEQUENCE FROM N.A.
Tissue=Embryo;
MEDLINE=94094332; PubMed=8269517;
Klodkin A.L., Matthes D.J., Goodran C.S.;
"The semaphorin genes encode a family of transmembrane and secreted
growth cone guidance molecules.";
Cell 75:1389-1399(1993).
11.
FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
SUBCELLULAR LOCATION: Type I membrane protein.
SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
SIMILARITY: Contains 1 sema domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
CC EMBL: L26980; AAA16609.1;
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin_repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01437; PSI; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
KW Signal; Developmental protein; Transmembrane; Glycoprotein;
KW Neurogenesis.
FT SIGNAL 1 20
FT CHAIN 21 712
FT DOMAIN 21 601
FT TRANSMEM 602 622
FT DOMAIN 623 712
FT DOMAIN 213 536
FT DOMAIN 473 473
FT DOMAIN 612 616
FT CARBHYD 42 42
FT CARBHYD 69 69
FT CARBHYD 161 161
FT CARBHYD 265 265
FT SEQUENCE 712 AA; 79751 MW; C5734FESB9C09FE3 CRC64;
Query Match 15.5%; Score 845; DB 1; Length 712;
Best Local Similarity 28.9%; Pred. No. 8e-47;
Matches 230; Conservative 130; Mismatches 270; Indels 166; Gaps 30;
QY 12 LHPFAGAGFEDESSE-----SIHGNTYKQYVVFVGHKPGKNTTORHLDIQVIMIN- 65
DB 12 LIALCHAMPDSSKLIHPKSVSKSPT-----GNATFPD-----FIVNQ 54
QY 66 --GTLYIARCHIYVDI-DTSHTEIYCSKLLWKSKQACVDTCKMKGRKDECHNEIK 122
DB 55 DETSLVGRNRYKDSIFTSERK-----GGRIFWSSDANGCLCAGKTDCCQNTIR 110
QY 123 VLLKXNDALFVCGTNAPNFCNFYKMTLEPFD-BFSGVAPCPYDAKHANVALFADGK 181
DB 111 LLYSSEPKLVI-CGNSYKP-CRTVAPKEGYLVEKEVEGIGLCYPNPEHNSTSVSYNGQ 170
QY 182 LYSATVTDPLADAVYRSLGESPTLRVVKDSKALKPVPVQAVYGVYVFFFRNAY 241
DB 171 LFSATVADFGSGDPLIYRE-----PQRTSLDGLKOLNAPNFVNSVAYGVYVFFFRNAY 225
QY 242 EYNTMGKVVPFPAQVQNVKGGSCRVLEKQWTFELKARLNCVFGDSHFYFNLQAVTD 302
DB 226 EYVNGCVIYRVARVCKDCKGGFHQSSDR-WTSLKARLNCVFGDSHFYFNLQAVTD 294
QY 302 VI--RINGRD---VVLATFSTPYNSIPGSVAVCAVDYLDIASVFTGRFKEQKSPDSTNPF 356
DB 285 IVEGRVNSDDSKLIYG-LITPVNAIGGSA-CAYQADILAVFEGSPKQETINSNKLVP 344
QY 357 PDERVPKPRPCCAGSSSRJERYATSNFPPDITPIKTHPLMDEAVPSIFRPFWRMYR 416
DB 345 PQNLVPEPRPGQGVPSR-----LPDKNVNFIKTHSLMED-VPAJFGKPLVLRVSL 395
QY 417 RYRLTKIANDTAGPYQNH--TVVFLGSEKGIILKLARLGNSEGLNDSLFLEKSNVNS 474
DB 396 QRFRTAITVDPQVKTINQYLVLVIGTDDGVK-----430
QY 475 EKCSYDGVEDKRMGMCLDRASSLYVAFSTCV-----IKVP-----511
DB 431 -----AVNIPKRHAKALLRYKRTSVHPHGPVQLKIAPGVGVVVGKDEI 478
QY 512 -----LGRCEHGRKCKKTCIASRDPYCGWIKEGGASHLSPNSRGT-F-EQTERGN-----561
DB 479 RLANNHCASKTEC-KDCVLEQDPHCAMDANKQLCVSIDTIVSYRFLICDVVVGDCNCK 537

```

562 ---TDLGSCHEPVALNGSSLLPSTTSSTAGCTESAGGLDMLKHLULSPDSTEP 618  
 538 SGTQKXVTKK-----PSEVENET-----NSIDKCL--DSST 572  
 619 LGAVSHHQKKGVIRESYKQ---HDCVPPVT---LALAVILAPWGA---VSGITV 670  
 573 LKLTGLDSDSDCPV-SENSIGCAVAGQOLVIVTAGTLHVVVAVSIVGLFWKXGLSV 631  
 671 YCV-CTHRKQVAVVQRE-KELTHSRGSM-SVTKLSGLFSTQKDKPKPAITP 725  
 632 FAXHSDSCVPEAFFTEQNHLELSANQCTGYLTPRANKAVMLVWVSSSTPRPKXNLD 691  
 726 LMHGKLCAPGNTAKY 74:  
 692 VSKDLIASDGTQK: 707

RESULT 15

Z1B BRARE  
 D S21B BRARE STANDARD; PRC: 778 AA.  
 C QW4686;  
 T 30-MAY-2000 (Rel. 39, Created:  
 T 30-MAY-2000 (Rel. 39, Last sequence update)  
 T 15-SEP-2003 (Rel. 42, Last annotation update)  
 E Semaphorin Z1B precursor (Semaphorin 1B) (Sema-21B).  
 N SEMA21B OR SEMA3A5  
 S Brachydanio rerio (Zebrafish) (Danio rerio).  
 C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 C Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 C Cyprinidae; Danio.  
 N NCBI\_TaxID=7955;  
 X [1]  
 P SEQUENCE FROM N.A.  
 X MEDLINE:99425174; PubMed:10495275;  
 X Roos M., Schachner M., Bernhardt R.R.,  
 T "Zebrafish semaphorin z1b inhibits growing motor axons in vivo".  
 T Mech. Dev. 87:103-117(1999).  
 C -!- FUNCTION: NIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE  
 C OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A  
 C REPUSIVE DOMAIN IN THE POSTERIOR SOMITE.  
 C -!- SUBCELLULAR LOCATION: Secreted (by similarity).  
 C -!- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND  
 C IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY  
 C VENTRALLY EXTENDING MOTOR AXONS.  
 C -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 C -!- SIMILARITY: Contains 1 sema domain.  
 C -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 C  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL outstation  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (See <http://www.isb.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C  
 R EXBL: AF063382; AA028101.1;  
 R ZFIN: ZDB-GENE-991209-6; semaphab.  
 R InterPro: IPR007110; Ig-like.  
 R InterPro: IPR003593; Ig.  
 R InterPro: IPR003006; Ig\_YHC.  
 R InterPro: IPR003659; plexin-like.  
 R InterPro: IPR002165; plexin\_repeat.  
 R InterPro: IPR001627; Sema.  
 R Pfam: PF00047; Ig.  
 R Pfam: PF01437; PSI.  
 R Pfam: PF01403; Sema.1.  
 R SMART: SM00409; IG.  
 R SMART: SM00423; PSI.  
 R SMART: SMC0630; Sema.1.  
 R PROSITE: PS00835; IGLIKE.  
 W Signal: Immunoglobulin domain; Multigene family; Neurogenesis;  
 W Developmental protein; Glycoprotein.

FT	SIGNAL	1	17	POTENTIAL
FT	CHAIN	18	778	SEMAPHORIN Z1B
FT	DOMAIN	24:	539	SEMA
FT	DOMAIN	579	668	IG-LIKE C2-TYPE
FT	DOMAIN	721	776	ARG/LYS-RICH (BASIC)
FT	DISULFID	652	716	BY SIMILARITY
FT	CARBOHYD	54	54	N-LINKED (GLCNAC) (POTENTIAL)
FT	CARBOHYD	127	127	N-LINKED (GLCNAC) (POTENTIAL)
FT	CARBOHYD	593	593	N-LINKED (GLCNAC) (POTENTIAL)
SC	SEQUENCE	778 AA;	59304 MW;	4036F432AE21995 CRC64;

Query Match: 15.58; Score 845; DE 1; Length 778;  
 Best Local Similarity 34.38; Pred. No. 9e-47;  
 Matches 194; Conservative 94; Mismatches 204; Indels 74; Gaps 17;

QY	66	STLYIARQHIYTVD-IDTSHTIEIYCKKLTKWKSRAQVDTCRMKXK-KDCENIKV	123
DB	68	GRLEFVGAKDHLV-SFNLYVDINMQQL-----SWPSPSPRRDECKKAGKQVKQKCANIKV	122
QY	124	LLKKNDDCALFVCGTNAKFNESCNYKM-----DTLEPEGDEP-SGVARCPYDAKHANVALF	177
DB	123	LQPNQTHLVACGTGAFHPVCAHVEVSKRSEDNTFRUGSSFENGGRGKSPYDPKQOTASML	182
QY	178	ADGKLYSATVTDFLAIDAVIYRSLGESPT-RTVKHDSKWLKPEYFV-----CAVDY	228
DB	183	IDGELYAGTSADFMGRDPAIFPTLGRKHPIRTEQHDSRWLNDPRFVSVHLIPESQNAED-	241
QY	229	GDYIYFPREIAVEYNNYKGVFFRVAVQVCKNDMGSGQRVLEKQWTKFLKARLNCSPVG-	287
DB	242	-DKYLFPRFENAIIDGEOISKATHARIQCLCKNDFFGG-HRSLVNKWTFLKARVCSVPGL	299
QY	288	---DSHEVENIQAQVTVQVIRINGRD---VVALTFSTPYNSIPGSAVCAVMLEIASVET	340
DB	300	NGIDTHP-----DELQVFLKSSKDPKNPIIYAVFTTSSNFKGSVAVCYSMADIRRVFL	354
QY	341	GRPKEQKSPDSTWTPVDPDERVFKPRPGCCAGSSSLERYATSNPEPDTLNIKTHPLMDE	400
DB	355	GPVAHROGPNVQWYFPLK-RVPYPRPGTGP-SKFDGPFESTKDPDDVITFARSHPAMYN	412
QY	421	AVPSIFRPPWELRTVYRILTKIANDVAAAGYQVHTVVFLOSEKGIILKFLARIGNSGFL	460
DB	413	PVFETNNHPIIKTCDVYQFTQIVVDVREAEADGGYDYNFCTDMGTQVJKVYSIPEGTHD	472
QY	481	NDSJFLEEMSYNSEKSYDGVDEKRMGMOLDRASSSLVVAESTCVIKVPLGRCERHGK	520
DB	473	LEVALLEEMTVFR-----EPTATCAMEJSTKQQQJYLSAIGVSQKPLHRCDDVYVK	523
QY	521	CKKTCIASRDYPYCGWIKEGGACSHLSPNS-RLTPEQDIERGNTDGLGCHNSFYALNGHS	579
DB	574	ACAECLAKDQYCAW--DSSQCSRYFPAKBRTRQDIAND-----	563
QY	580	SSLIPSTTTSSTAGQGVESRGGMLD 605	
DB	564	---PLTCCSLQCHDEADAGEAGLID 585	

Search completed: October 23, 2003, 17:09:37  
 Job time : 22 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 CompuGen Ltd.  
 X protein - protein search, using sw -mode  
 Run on: October 23, 2003, 17:09:07 : Search time 37.2 seconds  
 (without alignments)  
 4653.056 Million cell updates/sec

US-09-856-681-2  
 Perfect score: 5450  
 Sequence: 1 XRSBALLYFTLLHFGAGF.....PKKPSFALSTSMKPNDACT 1030

Scoring table: BLOSUM62  
 Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 255252604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

## Database :

SPTREMBL.23:\*  
 1: sp\_archea.\*  
 2: sp\_bacteria.\*  
 3: sp\_fungi.\*  
 4: sp\_human.\*  
 5: sp\_invertebrate.\*  
 6: sp\_mammal.\*  
 7: sp\_mhc.\*  
 8: sp\_organole.\*  
 9: sp\_plant.\*  
 10: sp\_protist.\*  
 11: sp\_rhodospirillum.\*  
 12: sp\_virus.\*  
 13: sp\_viruses.\*  
 14: sp\_unclassified.\*  
 15: sp\_virus.\*  
 16: sp\_bacteriophage.\*  
 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5450	100.0	1030	4	Q9H2E6
2	5431.5	99.7	1049	4	Q9P2H9
3	5024.5	92.2	1005	11	Q9EQ71
4	3677	67.5	699	4	Q965W4
5	2980.5	54.7	874	4	Q965W8
6	2963	54.4	887	11	Q8BUT0
7	2961	54.3	605	11	Q8BXZ7
8	2952	54.2	562	4	Q965W4
9	2947	54.1	562	4	Q8NC49
10	2621.5	48.1	507	4	Q965W4
11	2368	43.4	1073	4	Q8N9V4
12	2349	43.1	1017	4	Q8N9V5
13	2315.5	42.5	998	4	Q8N9V6
14	2314	42.5	1011	4	Q8N9V3
15	2314	42.5	1022	4	Q9P249
16	2213	40.6	418	4	Q965U9

## ALIGNMENTS

## RESULT :

Q9H2E6 PRELIMINARY: PRT: 1030 AA.  
 ID Q9H2E6  
 AC Q9H2E6  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Semaphorin SEMA6A.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20564339; PubMed=1093894;  
 RA Klostermann A., Lutz S., Gertler P., Bell G.  
 RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the Enabled/Vasodilator-stimulated Phosphoprotein-like Protein (EVL) via a novel carboxyl-terminal Tyrosin-like domain."  
 RL J. Biol. Chem. 275:39647-39653 (2000).  
 DR EMBL; AF279656; AAC29378.1;  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF01403; Sema; 1.  
 DR SMART; SM00423; PSI; 1.  
 DR SMART; SM00630; Sema; 1.  
 SC SEQUENCE 1030 AA; 114368 MW; A57E79C10AEC4B34 CRC64;  
 Query Match 100.0%; Score 5450; DB 4; Length 1030;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1030; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 XRSBALLYFTLLHFGAGFPEDEPISISGHGNYTKQVGVGHKPGKNTTQRHPIQDM 60  
 DB 1 XRSBALLYFTLLHFGAGFPEDEPISISGHGNYTKQVGVGHKPGKNTTQRHPIQDM 60  
 QY 61 IMWNGTLYAAROHITVDITDTSHTTEIYCSKLTWKSRQADVDTCRMKGKHKDECHNF 120  
 DB 61 IMWNGTLYAAROHITVDITDTSHTTEIYCSKLTWKSRQADVDTCRMKGKHKDECHNF 120



y	121	IKVLLKKNDALFVCGTNAENPSCBNYKMDLBPFGDEFSQACSCPYDAKHANVALFADG	130
b	121	IKVLLKKNDALFVCGTNAENPSCBNYKMDLBPFGDEFSQACSCPYDAKHANVALFADG	130
y	181	KLYSATVTDALDAVAYRSLGSSPTLRITKHSKWLKEPYFQAVDYGDIYFFPREIA	240
b	181	KLYSATVTDALDAVAYRSLGSSPTLRITKHSKWLKEPYFQAVDYGDIYFFPREIA	240
y	241	VEYNTMGVFFRVAQCKNDMGSSQRYLSEKQWTSFLKAPLNGSVGDSHFYFNILQAVT	300
b	241	VEYNTMGVFFRVAQCKNDMGSSQRYLSEKQWTSFLKAPLNGSVGDSHFYFNILQAVT	300
y	301	DVIRINGRDVVLATFSTPYKSIIPSAVCAVMDIASVYTRKEKQKPDSTWTFVPCR	360
b	301	DVIRINGRDVVLATFSTPYKSIIPSAVCAVMDIASVYTRKEKQKPDSTWTFVPCR	360
y	361	VPKPRGCCAGSSSLERYATSNBFPDQTLNFKTHPLMDNAVPSIFNRPWFLRTWVYEL	420
b	361	VPKPRGCCAGSSSLERYATSNBFPDQTLNFKTHPLMDNAVPSIFNRPWFLRTWVYEL	420
y	421	TKIAVDTAAGFYQNHVYVFLGSEKGIILKPLARIGNSGFLNDSFLSEMSVYNSEKSYD	480
b	421	TKIAVDTAAGFYQNHVYVFLGSEKGIILKPLARIGNSGFLNDSFLSEMSVYNSEKSYD	480
y	481	GVEDKRMGMQLDRASSSLYVAFSTCVIKVPLGRCEHKGKCTCIASRDYCGWIKEGG	540
b	481	GVEDKRMGMQLDRASSSLYVAFSTCVIKVPLGRCEHKGKCTCIASRDYCGWIKEGG	540
y	541	ACSHLSNPNRUTFECDIERGNTGCGCHNSFVALNDSTPLPQNEYSYNTVYGHSSSL	600
b	541	ACSHLSNPNRUTFECDIERGNTGCGCHNSFVALNDSTPLPQNEYSYNTVYGHSSSL	600
y	601	GGMLDWKHLSDPSDTPLGAVSHNQKKGVIRESYLKGSQVPTLLAVALIAEV	660
b	601	GGMLDWKHLSDPSDTPLGAVSHNQKKGVIRESYLKGSQVPTLLAVALIAEV	660
y	661	MGAVFSGITVYVCDHPRKQVAVYQKEKELTHSPFGSYSSVTKLSGLEQDTSQPKIE	720
b	661	MGAVFSGITVYVCDHPRKQVAVYQKEKELTHSPFGSYSSVTKLSGLEQDTSQPKIE	720
y	721	ALITPLMHNKSLATPGTAKMLIQAQHHLDLTAIPFESTPTLQCKRPFSSRSEKRN	780
b	721	ALITPLMHNKSLATPGTAKMLIQAQHHLDLTAIPFESTPTLQCKRPFSSRSEKRN	780
y	781	ONLINACTKQPPMGSPVITDPLRASPSHISVYVLPITQSGYQHYVQDQPMGEVAQ	840
b	781	ONLINACTKQPPMGSPVITDPLRASPSHISVYVLPITQSGYQHYVQDQPMGEVAQ	840
y	841	VALEDOAATLEYXTIKEHLSKSPNHGWNVENLDSLPPKVPORBASLGPPGASTQTL	900
b	841	VALEDOAATLEYXTIKEHLSKSPNHGWNVENLDSLPPKVPORBASLGPPGASTQTL	900
y	901	SKLEMHSSSYGVYKRYPTNLSLRHQATLKNNTNNSKSSLSRNSQSGRCNPP	960
b	901	SKLEMHSSSYGVYKRYPTNLSLRHQATLKNNTNNSKSSLSRNSQSGRCNPP	960
y	961	PACRVDSICVHSSQPSQAVTVSQPSLNAVNSLTRSLGRTPSLKQVPPKPSFAPLS	1020
b	961	PACRVDSICVHSSQPSQAVTVSQPSLNAVNSLTRSLGRTPSLKQVPPKPSFAPLS	1020
y	1021	TSMKPNDACT 1030	
b	1021	TSMKPNDACT 1030	

RESULT 2  
9P2H9  
D Q9P2H9 PRELIMINARY; PRT: 1049 AA.  
C Q9P2H9  
T 01-OCT-2000 (TrEMBLrel. 15, Created)  
T 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
T 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE	Hypothetical protein KIAA1368 (Fragment).	
GN	KIAA1368.	
CS	Homo sapiens (Human).	
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
CC	NCBI_TaxID=9606;	
RI	11_1	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Brain;	
RX	MEDLINE=20181126; PubMed=10718198;	
RA	Magase T., Kikuno R., Ishikawa K., Hirosewa Y., Obara C.,	
RT	Prediction of the coding sequences of unidentified human genes.XV.	
RT	The complete sequences of 150 new cDNA clones from brain which code	
RT	for large proteins in vitro.	
RU	DNA Res. 7:65-73(2000).	
DR	EMBL; AB017789; BAA92606.1; -	
DR	Genbank; HGNC:10738; SEMA6.	
DR	InterPro; IPR003659; Plectin-like.	
DR	InterPro; IPR003627; Sema.	
DR	Pfam; PF01403; Sema; 1	
DR	SMART; SM0423; PSI; 1	
DR	SMART; SM0630; Sema; 1	
KW	Hypothetical protein.	
FT	NON_TER 1	
SC	SEQUENCE 1049 AA; 11651; MW: 778120AOC7A8A8A CRC64;	
Query Match	99.7%; Score 5431.5; DB 4; Length 1049;	
Best Local Similarity	98.4%; Pred. No. 0;	
Matches 1010; Conservative C; Mismatches 0; Indels 17; Gaps 1;		
OY	1 MPSEALLLYFTLLFAGAGFEDSEPISSHGNTKCYPVGVGHKPGPNTTCRRLDIOM	60
OB	3 KRSEALLLYFTLLFAGAGFEDSEPISSHGNTKCYPVGVGHKPGPNTTCRRLDIOM	62
OY	61 IVIMNGTLYIAAROHIVTVDTSHTEELYSKSLTWKSPQADYDTCPYKGRKDFCHNF	120
OB	63 IVIMNGTLYIAAROHIVTVDTSHTEELYSKSLTWKSRQADYDTCPYKGRKDFCHNF	122
OY	121 KULLKKNDALFVCGTNAENPSCBNYKMDLBPFGDEFSQACSCPYDAKHANVALFADG	180
DE	123 KULLKKNDALFVCGTNAENPSCBNYKMDLBPFGDEFSQACSCPYDAKHANVALFADG	182
OY	181 KLYSATVTDALDAVAYRSLGSSPTLRITKHSKWLKEPYFQAVDYGDIYFFPREIA	240
OB	183 KLYSATVTDALDAVAYRSLGSSPTLRITKHSKWLKEPYFQAVDYGDIYFFPREIA	242
OY	241 VEYNTMGVFFRVAQCKNDMGSSQRYLSEKQWTSFLKAPLNGSVGDSHFYFNILQAVT	300
OB	243 VEYNTMGVFFRVAQCKNDMGSSQRYLSEKQWTSFLKAPLNGSVGDSHFYFNILQAVT	302
OY	301 DVIRINGRDVVLATFSTPYKSIIPSAVCAVMDIASVYTRKEKQKPDSTWTFVPCR	360
OB	303 DVIRINGRDVVLATFSTPYKSIIPSAVCAVMDIASVYTRKEKQKPDSTWTFVPCR	362
OY	361 VPKPRGCCAGSSSLERYATSNBFPDQTLNFKTHPLMDNAVPSIFNRPWFLRTWVYEL	420
OB	363 VPKPRGCCAGSSSLERYATSNBFPDQTLNFKTHPLMDNAVPSIFNRPWFLRTWVYEL	422
OY	421 TKIAVDTAAGFYQNHVYVFLGSEKGIILKPLARIGNSGFLNDSFLSEMSVYNSEKSYD	480
OB	423 TKIAVDTAAGFYQNHVYVFLGSEKGIILKPLARIGNSGFLNDSFLSEMSVYNSEKSYD	482
OY	481 GVEDKRMGMQLDRASSSLYVAFSTCVIKVPLGRCEHKGKCTCIASRDYCGWIKEGG	540
OB	483 GVEDKRMGMQLDRASSSLYVAFSTCVIKVPLGRCEHKGKCTCIASRDYCGWIKEGG	542
OY	541 ACSHLSNPNRUTFECDIERGNTGCGCHNSFVALNDSTPLPQNEYSYNTVYGHSSSL	563
OB	543 ACSHLSNPNRUTFECDIERGNTGCGCHNSFVALNDSTPLPQNEYSYNTVYGHSSSL	602
OY	581 PSTTTSDTAQGYESGGMLDWKHLSDPSDTPLGAVSHNQKKGVIRESYLKGD	643
OB	603 PSTTTSDTAQGYESGGMLDWKHLSDPSDTPLGAVSHNQKKGVIRESYLKGD	662





```

482 NTNSNSSILSRKQSGEDNPPAPQRCVSTQVSSOSGQAVYVSQPSLNAVNSLTR 541
998 SGLKTPSSJKPVPPKPSFAPJSTSMKPKDACT 1030
542 SGLKTPSSJKPVPPKPSFAPJSTSMKPKDACT 574

RESULT 6
88BUC0
D CSBUC PRELIMINARY: PRT: 587 AA.
C C9BUC;
T 01-MAR-2003 (TREMBlrel. 23, Created);
T 01-MAR-2003 (TREMBlrel. 23, Last sequence update);
T 01-MAR-2003 (TREMBlrel. 23, Last annotation update);
E Sema domain.
S Mus musculus (Mouse).
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
X NCBI_TaxID:10090;
W
SEQUENCE FROM N.A.
P STRAIN=C57BL/6J; TISSUE=Cerebellum;
X MEDLINE=22354683; PubMed=12466851;
A The RIKEN Consortium.
A "Analysis of the mouse transcriptome based on functional annotation of
T 60,770 full-length cDNAs."
R Nature 420:563-573(2002);
J EMBL; AK042751; BAC31352.1;
Q SEQUENCE 587 AA; 66044 MW; 8E9C7B102C4DB97A CRC64;

```

```

Query Match 54.4%; Score 2963; DB 11; Length 587;
Best Local Similarity 95.1%; Pred. No. 1.8e-232;
Matches 549; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

1 MRSEALLLYTLHFAGAGFPDESEPISSHGNYTKQYPVFGKPGGRNTTQRRLDIOM 60
1 MPAALLLCITLHCAGAGFPDESEPISSHGNYTKQYPVFGKPGGRNTTQRRLDIOM 60
61 IMXNGTLYIAARCHIYTVDTISHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
61 IMXNRTLYVAARDHIYTVDTISHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
121 IKVLLKNDALFVCGTNAPNSCRNYKMTLEPPGDEFSGMARCPYDAKHANVAFAG 180
121 IKVLLKNDCTLFVCGTNAPNSCRNYKMTLEPPGDEFSGMARCPYDAKHANVAFAG 180
181 KLYSATVDTFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYCAVDYGYIYFFPREIA 240
181 KLYSATVDTFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYCAVDYGYIYFFPREIA 240
241 VEYNTMGKVPFPAVACVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
241 VEYNTMGKVPFPAVACVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
301 DVIRINGRDVVLATFTSTPKNSIPGSAVCAYDMLEIANVFGRFKEQKSPDSTWTPVEDR 360
301 DVIRINGRDVVLATFTSTPKNSIPGSAVCAYDMLEIANVFGRFKEQKSPDSTWTPVEDR 360
361 VPKPRPGCCAGSSSLEKYATSNPEPDDTLNFITKTHPLMDEAVPSINRPWFRTMYRYRL 420
361 VPKPRPGCCAGSSSLEKYATSNPEPDDTLNFITKTHPLMDEAVPSINRPWFRTMYRYRL 420
421 TKIADVTAAGPYCNHTVFWFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVYNSKCSYD 480
421 TKIADVTAAGPYCNHTVFWFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVYNSKCSYD 480
481 GVEDKRIYNGQIDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWIKEGG 540
481 GVEDKRIYNGQIDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWIKEGG 540
491 GVEDKRIYNGQIDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWIKEGG 540
541 ACSHLSPNSRLTFEQDIERGNTDGLGCHNSFVALNG 577

```

```

541 SCAHLSPNSRLTFEQDIERGNTDGLGCHNSFVALNG 577
PRELIMINARY: PRT: 605 AA.
C C9EXZ7;
T 01-MAR-2003 (TREMBlrel. 23, Created);
T 01-MAR-2003 (TREMBlrel. 23, Last sequence update);
T 01-MAR-2003 (TREMBlrel. 23, Last annotation update);
E Sema domain.
S Mus musculus (Mouse).
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
X NCBI_TaxID:10090;
W
SEQUENCE FROM N.A.
P STRAIN=C57BL/6J; TISSUE=Cerebellum;
X MEDLINE=22354683; PubMed=12466851;
A The RIKEN Consortium.
A "Analysis of the mouse transcriptome based on functional annotation of
T 60,770 full-length cDNAs."
R Nature 420:563-573(2002);
J EMBL; AK042751; BAC31352.1;
Q SEQUENCE 605 AA; 62299 MW; B4FE713BE8AF6E90 CRC64;

Query Match 54.3%; Score 2961; DB 11; Length 605;
Best Local Similarity 94.5%; Pred. No. 2.8e-232;
Matches 550; Conservative 14; Mismatches 18; Indels 3; Gaps 0;

1 MRSEALLLYTLHFAGAGFPDESEPISSHGNYTKQYPVFGKPGGRNTTQRRLDIOM 60
1 MPAALLLCITLHCAGAGFPDESEPISSHGNYTKQYPVFGKPGGRNTTQRRLDIOM 60
61 IMXNGTLYIAARCHIYTVDTISHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
61 IMXNRTLYVAARDHIYTVDTISHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
121 IKVLLKNDALFVCGTNAPNSCRNYKMTLEPPGDEFSGMARCPYDAKHANVAFAG 180
121 IKVLLKNDCTLFVCGTNAPNSCRNYKMTLEPPGDEFSGMARCPYDAKHANVAFAG 180
181 KLYSATVDTFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYCAVDYGYIYFFPREIA 240
181 KLYSATVDTFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYCAVDYGYIYFFPREIA 240
241 VEYNTMGKVPFPAVACVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
241 VEYNTMGKVPFPAVACVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
301 DVIRINGRDVVLATFTSTPKNSIPGSAVCAYDMLEIANVFGRFKEQKSPDSTWTPVEDR 360
301 DVIRINGRDVVLATFTSTPKNSIPGSAVCAYDMLEIANVFGRFKEQKSPDSTWTPVEDR 360
361 VPKPRPGCCAGSSSLEKYATSNPEPDDTLNFITKTHPLMDEAVPSINRPWFRTMYRYRL 420
361 VPKPRPGCCAGSSSLEKYATSNPEPDDTLNFITKTHPLMDEAVPSINRPWFRTMYRYRL 420
421 TKIADVTAAGPYCNHTVFWFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVYNSKCSYD 480
421 TKIADVTAAGPYCNHTVFWFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVYNSKCSYD 480
481 GVEDKRIYNGQIDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWIKEGG 540
481 GVEDKRIYNGQIDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWIKEGG 540
541 ACSHLSPNSRLTFEQDIERGNTDGLGCHNSFVALNG 582
541 SCAHLSPNSRLTFEQDIERGNTDGLGCHNSFVALNDISTP 582

```

## RESULT 8

```

96SY4
D Q96SY4 PRELIMINARY: PRT: 562 AA.
C Q96SY4:
T 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
T 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
T 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
E Hypothetical protein FJ14565.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
X NCBI_TaxID=9606;
Y .
F SEQUENCE FROM N.A.
A Isogai T., Ota T., Nishikawa T., Sugiyama T., Otsuki T., Suzuki Y.,
A Nishikawa T., Nagai K., Sugano S., Ishii S., Kawai H., Sugiyama T.,
A Watanabe S., Kimura K., Nakamura Y., Kojima S., Nagahara K.,
A Watanabe S., Kimura K., Nakamura Y., Kojima S., Nagahara K.,
A Yamamoto C., Kimura K., Nakamura Y., Kojima S., Nagahara K.,
A Kinomiya K., Iwayanagi T.,
T "NEDO human cDNA sequencing project.";
R Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
R EMBL: AK029471; BA355136;
R InterPro: IPR003639; Flexin-like.
R SMART: SMC423; PSI;
R Hypothetical protein.
W SEQUENCE 562 AA; 61313 MW; 6AB3695FAD1DD78A CR664;

Query Match 54.2%; Score 2952; DB 4; Length 562;
Best Local Similarity 102.0%; Pred. No. 136-231;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

469 MSYNEKSYGVDEKRIKNGVQLDRASSSLVAFSTCVIKVPLGRCERHGKCKKTCIAS 529
: MSYNEKSYGVDEKRIKNGVQLDRASSSLVAFSTCVIKVPLGRCERHGKCKKTCIAS 60
529 RDPYCKWIKEGGACSHLSNSRLTFQD-ERGNLTLGDCRHSFVALNCHSSSLPSTTT 589
: RDPYCKWIKEGGACSHLSNSRLTFQD-ERGNLTLGDCRHSFVALNCHSSSLPSTTT 120
61 RDPYCKWIKEGGACSHLSNSRLTFQD-ERGNLTLGDCRHSFVALNCHSSSLPSTTT 120
589 SDSTAEGYESRGKGLDWKHELDSPSTPLGAVSHHCKKGVIRESYLKHQDLVPV 649
121 SDSTAEGYESRGKGLDWKHELDSPSTPLGAVSHHCKKGVIRESYLKHQDLVPV 180
649 TLLAIAVILAFVWGAVFSGITVYVCDHRKQVAVVCRKEKELTHSRGSMSSVTKLSG 709
181 TLLAIAVILAFVWGAVFSGITVYVCDHRKQVAVVCRKEKELTHSRGSMSSVTKLSG 240
709 FGDTQSKDPKEAILELPLMNGKLTATPGNTAKMLIKADQHHLDLTALPTPESTPTLQCKR 769
241 FGDTQSKDPKEAILELPLMNGKLTATPGNTAKMLIKADQHHLDLTALPTPESTPTLQCKR 300
769 KPSRGSEKERNONLNACTKQFNGSPVPTDPLRASPSHIPSVVLPITQCGYCHE 829
301 KPSRGSEKERNONLNACTKQFNGSPVPTDPLRASPSHIPSVVLPITQCGYCHE 360
829 YVDQPKXSEVAQMALEDQAATLEYKTKHELSKSPNHNVENLDSLPKVPQVPEASL 889
361 YVDQPKXSEVAQMALEDQAATLEYKTKHELSKSPNHNVENLDSLPKVPQVPEASL 420
889 GPPGASLSQGLSKRLMHSSSYGVYKSYPTNSLTSHQATTLKRNNTNSNSHLS 949
421 GPPGASLSQGLSKRLMHSSSYGVYKSYPTNSLTSHQATTLKRNNTNSNSHLS 480
949 RNSGFGGNPPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTFSLKP 1009
481 RNSGFGGNPPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTFSLKP 540
1009 DVPPKPSFAPLSTSMKPNDACT 1030
541 DVPPKPSFAPLSTSMKPNDACT 562

```

RESULT 10  
Q96SY4

```
Q96T04 PRELIMINARY; PRT: 507 AA.
Q96T04
01-DEC-2001 (TrEMBLrel. 19, Created:
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E Hypothetical protein F5J4533.
S Homo sapiens (human)
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
P SEQUENCE FROM N.A.
A Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuka T., Suzuki Y.,
A Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
A Watanabe S., Hosokawa T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
A Takahashi Y., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
A Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
A Yamamoto S., Wakamatsu A., Nakamura Y., Nagahara K., Masuko Y.,
A Ninomiya K., Iwayanagi T.,
I "NEDO human cDNA sequencing project.";
J Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
R EMBL; AK027439; BAB55111.1;
R InterPro; IPR003659; Plectin-like.
R SMART; SM00423; PS2.1;
M Hypothetical protein.
Q SEQUENCE 507 AA; 55464 MW; 8CC567B438C51B39 CRC64;

Query Match 48.1%; Score 2621.5; DB 4; Length 507;
Best Local Similarity 90.2%; Pred. NO. 9.9e-205;
Matches 507; Conservative C; Mismatches C; Indels 55; Gaps 1;

469 XSVYNSEKCSYGVQEKRMGMQDASSLYAFSTCVKPLGRCHGKCKTCIAS 528
1 XSVYNSEKCSYGVQEKRMGMQDASSLYAFSTCVKPLGRCHGKCKTCIAS 60
529 RDPYCGMKEGGACSHLSPNSRIFPDQIERGNTGLGCHNSFVALNGHSSSLPSITT 588
61 RDPYCGMKEGGACSHLSPNSRIFPDQIERGNTGLGCHNSFVALNGHSSSLPSITT 128
589 SDSTAQCGYSRGMNDWKHLDSFSTDFLGAVSSHNLKKGVPESYIKOHQLQPV 648
109 -----GVRESYKGRQLQPV 125
649 TLLAIAVILAFVNGAVFSGITVYVCDHRPKQVAVYCRKEKLUHSPGYSYVKLSGL 708
126 TLLAIAVILAFVNGAVFSGITVYVCDHRPKQVAVYCRKEKLUHSPGYSYVKLSGL 135
709 FGDQSKDPKPRALITFLMHKGLATGNTAKMLKADQCHHLQDLALPTFSTPLQQR 768
126 FGDQSKDPKPRALITFLMHKGLATGNTAKMLKADQCHHLQDLALPTFSTPLQQR 745
769 KPSRGSPEKERNQNLNACTKDYIPMSGVIPDLPILASPSHFYVWLPITQGYQHE 824
246 KPSRGSPEKERNQNLNACTKDYIPMSGVIPDLPILASPSHFYVWLPITQGYQHE 305
829 YVDCPKMSEVAQMALEDCAATLKYTKIKELHLSKSPNHWNLVENVLDSIFPKVFGREASL 888
306 YVDCPKMSEVAQMALEDCAATLKYTKIKELHLSKSPNHWNLVENVLDSIFPKVFGREASL 365
889 GPPGASLSTGLSKRLMHSSSYGYCYKASYVTNLSLTHSHQATLUKRNNTSSNSHLS 948
366 GPPGASLSTGLSKRLMHSSSYGYCYKASYVTNLSLTHSHQATLUKRNNTSSNSHLS 425
949 RNQSGFGSDPPAPQVQVSGVSGQSVQAVTVSSQSLKAVNSLTPSGSKRTPSLKP 1009
426 RNQSGFGSDPPAPQVQVSGVSGQSVQAVTVSSQSLKAVNSLTPSGSKRTPSLKP 465
1309 DVPPKPSFAFLSTSYKENDACT 1030
486 DVPPKPSFAFLSTSYKENDACT 507
```

RESULT 11

```
Q8NFY4 PRELIMINARY; PRT: 1073 AA.
Q8NFY4
01-OCT-2002 (TrEMBLrel. 22, Created:
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 6D isoform 4.
CS SEMA6B.
CS Homo sapiens (human)
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
P SEQUENCE FROM N.A.
R RESSUE=Brain;
RA CC X., Zhai Y., Wei H., Yu Y., Tang F., He F.;
RT "Homo sapiens semaphorin 6D isoform 4 (SEMA6D.4) mRNA, complete cds.";
BL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
EP EMBL; AF389429; AAY69452.1;
EP InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00630; Sema; 1.
Q SEQUENCE 1073 AA; 119672 MW; 7DCE4DFC5BF7CF9E CRC64;

Query Match 43.4%; Score 2368; DB 4; Length 1073;
Best Local Similarity 46.0%; Pred. NO. 1.5e-183;
Matches 512; Conservative 159; Mismatches 368; Indels 134; Gaps 30;

QY 1 YRSEALLLYFTLL---HFAGAGPEPDEPTISISGHNVTKQYVVFVGHKPGRNITQHRRLD 57
DY 1 XVPFLLCAYILLNVSCLRAVSPEDDEPLNTVDYHSQYVPVFRG-RPSGNEQ-HRLD 58
QY 58 IQMIMNMGTLYIAARDRIYTDISTHTBEIYCSKLTWKSQAQVDTCRMKGKHKDEC 117
DY 59 FQLMKIRLDLYIAGRDQVYTNLNMKPTETVFNKLTWRSQQRENCAMKGRHKDEC 118
QY 118 HNFVKVLKKNDDALFVCGTNAFNPSCRNVMDTLEPFGDFESGMARCPYDAKHAVLFP 177
DY 119 HNFVKVFPNDENWFVCGTNAFNPSCRNVYELSTLEYDGEESGLARCFDARQTVALLF 178
QY 178 ADGKLYSATVTDLAIDAVTYRSLGESPTLRTVYKDSKWLKEPYFVCAVDYGEYVFFFR 237
DY 179 ADGKLYSATVADFLASDAVYRSMGDSALRTIKYDSKWLKEPFLHATEYGNVYFFFR 236
QY 238 EIAVENTMKVYFPPVAQVCKNDMGSSORVLEKQWTSFLKALNCSVPGDSEHFENIQ 297
DY 239 EIAVEHNKQATSVRARCRKNDMGSSORVLEKQWTSFLKALNCSVPGDSEHFENIQ 298
QY 298 AVTDVIRINGRDVVLATFSTFYKSIPOSAVCAVMDIASVFTGRFEKQSPDSTWTPVP 257
DY 299 SITDIQNGIPTVGVPTTCLNSIFGSVAVCAFSMDIEKVPKGRFEKQTPDSVWTAVP 358
QY 358 DERVPKPRPCCAGSSSLERYATSNFPDPTLNFIKTHLMDEAVPSIENRWFLEMTVP 417
DY 359 ECKVPKPRPCCAGKAGHIAEAYKTSIDPDDETLSFKSHPLMDSAVPPIADEPMFTKTRV 418
QY 418 YSLTKIAVDTAAGPYQNHVYVJGSEKGLIKFLARIGSGGF-LNDSLPLEFMSVYNSK 476
DY 419 YSLTKIAVDSASAPYQNYVIYVGSAGNVLKVLAK--TSPSLNDSVLEELIENHAK 476
QY 477 CSYDGVDEKRMQMDRASSSYVAFSTCVIKVPLGRCHGKCKTCIASRDPYCGMI 536
DY 477 CSAENEEDKVKVSLQ-DKDIHIALYVAFSSCIIRIPLSRRCERYGSKKCIASRDPYCGMI 536
QY 537 KEGGACSHLSPNSRLT-FEODIERGNTGLGCHNSFVALNGHSSSLPSTTSSSTAGE 595
DY 537 SQ-GSGGRVTFGMLAEYEQTEFGNTAHLGQCH-----ELPTSTTPD----- 579
QY 596 GYBSRG-----MCD-MKILL-----DSP-----DSTSP 618
DY 580 -YKIFGPTSCYEVSSSVTTMASIFETPKVIDTWPKJTSRKPFVVDNPTSDFTTP 638
QY 619 LGAVSSHNHCKKGVIRESYLKGHDQLVPVTLIAIAVILAFVNGAVFSGTIVYVCD-HR 677
```





```

1 NCBI_TaxID=9606;
2
3 SEQUENCE FROM N.A.
4
5 TISSUE=Brain;
6
7 Qu X., Zhai Y., Wei H., Yu Y., Tang F., He F.;
8 "Homo sapiens semaphorin 6D isoform 2 (SEMA6D.2): mRNA, complete cds."
9 Submitted (JUN-2001) to the EMBL/GenBank/INSDC databases.
10
11 EMBL AF389430.1; AF389430.1;
12
13 InterPro: IPR001627; Pfam: PF01403; SMART: SM00630;
14
15 SEQUENCE 998 AA, 111730 MW, 38466672ERD5344 GRC64;
16
17 Query Match 42.5%; Score 2315.5; DB 4; Length 998;
18 Best Local Similarity 45.4%; Pred. No. 3,5e-179;
19 Matches 494; Conservative 155; Mismatches 292; Indels 148; Gaps 26;
20
21 1 MRSEALLVFTLL--HFAAGAPFDESEPISSHGNYTKQYVVFVGHKFGENTTQRRLD 57
22 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 1 MRVELLCAYLLKXVUSO.PAVSFPDEDEPDKTVDEYSRQYTVFAG-RPSGNEQ-HRLD 58
24 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 58 IQVIMMNGTVAARDHYTVVDIDTSHTEEVYCSKKLTWKSQRQADVTCRMKGKHQSC 117
26 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 59 FQMLKIRDTLYIAGRDQVTVNLEVPKTEFNKKLTWRSQQDRENCAMKGKHQSC 118
28 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
29 118 HNFIVKLVKKQDALLVCGTNAPNPSCRNYKYDTLEPPFGDEPSGKARCPCYDAKHANVAF 177
30 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 119 HNFIVKLVFERNDEWEVCGTNAPNPSCRNYKYDTLEPPFGDEPSGKARCPCYDAKHANVAF 178
32 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 178 ADGKLYSATVDFLAIDAVIYSSGSEPTDRTYKHSKWKKEFFYFVQAVYGDYIYFFR 237
34 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35 179 ADGKLYSATVDFLASDAVIYSSGSGALRTIKYDSKWIKREPHFHAIEYGVNVPFFR 238
36 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
37 238 EYAVENTNGKTVFPRVAGVCGNCGSORVLEKQWTFELARLNCVSGSHFENILQ 297
38 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 239 EYAVENHNGKAWERVARICNDVGGSORVLEKQWTFELARLNCVSGSHFENILQ 298
40 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 298 AVTDVIRINGRDVYATFSTPNIPGSVAVCYMDLQIASVFTGRFKESKPSDSTWVP 357
42 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 299 SITDIQINGIPTWGVFTQLNSIPGSVAVCAFSMDIEKVFKEQKTPDSVMTAVP 358
44 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
45 358 DERVPKPRGCCAGSSSLERYATSNFPDCTLNFIKTHPMDENVPSINRPFRLTMR 417
46 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
47 359 EDKVPKPRGCCAGKGLAEVATSDPDETLSEFKSHPLXDSAVPEIADPEWPKTEVP 418
48 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 418 YKTKAVDTNACPVNHWVFGSGEKGIIKELAPGNSGFINKSLFLEMSVYNSEK 476
50 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 419 YKTLAISVDHSAGPVQNYTVIFVSGEAGVVKVJAK--TSFPGKDSVLEEBEINYNK 476
52 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
53 477 CSYDGVEDKPIVGMQDRASSLVYAFSTCVIKVFLGRCERHCKKKTCTASHDFYQWI 536
54 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
55 477 CSAENSEDKKVLISLODKQHRLYVAFSSCIHRIPLGCEHYGSKKSCIASRDFYQWI 536
56 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 537 KEGGACSHLSPNSRLT-FEQDLERNTDGLGCHNSFVALNHSRSLSPSTTSSTAGH 595
58 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59 537 SC-GSGCRPTFGMLAEGYQCTEFKTHLGGCHG----- 570
60 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 596 GYASRSGKLDWKLHLLDPSDIDPLJAVSSHNCKGKIFESYLKGDOLYVETLLAV 655
62 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 571 -----VWKE--VSGSGS----- 593
64 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
65 656 ILAFVXGAVFSGITVYVCGD-HRRKDVAVYVORKEKETHSRAGSMESVTKSGLFG 710
66 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 594 PFAFVVGAFIAGVAVYCYRDVFKVQRK--HKDAESAQSCDSSGSKAKGLNGLPDSVK 651
68 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69 711 DTGSKDPKPEALITPLMNGKILATPNTANNK--KADQHLQDUTALPTPESTPLQOKRP 770
70 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 652 EYQCNDCSKLNSLLTSKELPPNGCTKSNVMDHRRGQPPLEALPTPESTPLVQLKTLQ 711
72 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 771 SRGSRWEMENQNLINACTKDYDPM3SPVLTDLPLRASPH-PSVTVLPITCCQYCHEY 829
74 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 712 AMKSHSEKAGH--GASRRETQFQFSSPPPHSPJ--SKGHIPSVIVPNATHYNTSFS 767

```

```

630 -----VDQF-----KYSEVAQVALEQDQATLEKVTKEHLSSKSPN----- 865
631 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 768 NSMAHKAERKLIQNDHPLTKSSKDHRRSVDSRNTL--NDLKLNDNNSPKAIMGDI 825
633 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
634 866 ---HGVNUNVENLDS---PPKVPQREASLGPFGASLSQTLGSKRLMHRHSSSY---GVY 916
635 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
636 926 QWAHQVULMDPMSSENVPPKVPNREASLYSPSTLPNSPKRYDVDFTPGVPMNTSLER 985
637 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
638 917 KSYFTNSLTRHQATTLKRNVTNSNSHJLNRNSGFRGSD-NPPAPGRVDSIQVHSSC 975
639 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
640 996 GCGYHKHSSQR-HS-SAMPK-NLNSFNGVLLSRQPSMRNGGYMPTTGAKVDYIO---- 938
641 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
642 996 PSCAVTVSRQPELAINYNLSI-----RSLKRTTSLKPEVPRKPSFAELSTSMKP 1025
643 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 999 --GTPSVHQLQPSLSQSSYFANGT-PRIGLRTPSLAKEDVPKPSFVPTSVTP 992
645 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
646
647 RESULT 14
648 QNPFY3
649 ID QNPFY3 PRESLININAPY; P87; 1011 AA.
650 AC QNPFY3;
651 CC 31-OCT-2002 (TREMblrel. 22, Created)
652 CC 31-OCT-2002 (TREMblrel. 22, Last sequence update)
653 CC 31-MAR-2003 (TREMblrel. 23, Last annotation update)
654 DE Semaphorin 6D isoform 1.
655 GN SEMA6D.
656 OS Homo sapiens (Human).
657 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
658 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
659 OX NCBI_TaxID=9606;
660 RN [-]
661 RE SEQUENCE FROM N.A.
662 RC T-SSUE=Brain;
663 RA Qu X., Zhai Y., Yu Y., Tang F., He F.;
664 RA "Homo sapiens semaphorin 6D isoform 1 (SEMA6D.1): mRNA, complete cds."
665 RL Submitted (JUN-2001) to the EMBL/GenBank/INSDC databases.
666 RL EMBL AF389430.1; AF389430.1;
667 CP InterPro: IPR001627; Pfam: PF01403; SMART: SM00630;
668 DR SMART: SM00630; Pfam: PF01403;
669 SC SEQUENCE 1011 AA, 115289 MW, 9068693633941B89 GRC64;
670
671 Query Match 42.5%; Score 2314; DB 4; Length 1011;
672 Best Local Similarity 45.4%; Pred. No. 3,5e-179;
673 Matches 494; Conservative 155; Mismatches 292; Indels 148; Gaps 26;
674
675 1 MRSEALLVFTLL--HFAAGAPFDESEPISSHGNYTKQYVVFVGHKFGENTTQRRLD 57
676 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
677 1 MRVELLCAYLLKXVUSO.PAVSFPDEDEPDKTVDEYSRQYTVFAG-RPSGNEQ-HRLD 58
678 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
679 58 IQVIMMNGTLYIAARDHYTVVDIDTSHTEEVYCSKKLTWKSQRQADVTCRMKGKHQSC 117
680 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
681 59 FQMLKIRDTLYIAGRDQVTVNLEVPKTEFNKKLTWRSQQDRENCAMKGKHQSC 118
682 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
683 118 HNFIVKLVKKQDALLVCGTNAPNPSCRNYKYDTLEPPFGDEPSGKARCPCYDAKHANVAF 177
684 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
685 119 HNFIVKLVFERNDEWEVCGTNAPNPSCRNYKYDTLEPPFGDEPSGKARCPCYDAKHANVAF 178
686 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
687 178 ADGKLYSATVDFLAIDAVIYSSGSEPTDRTYKHSKWKKEFFYFVQAVYGDYIYFFR 237
688 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
689 179 ADGKLYSATVDFLASDAVIYSSGSGALRTIKYDSKWIKREPHFHAIEYGVNVPFFR 238
690 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
691 238 EYAVENTNGKTVFPRVAGVCGNCGSORVLEKQWTFELARLNCVSGSHFENILQ 297
692 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
693 239 EYAVENHNGKAWERVARICNDVGGSORVLEKQWTFELARLNCVSGSHFENILQ 298
694 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
695 298 AVTDVIRINGRDVYATFSTPNIPGSVAVCYMDLQIASVFTGRFKESKPSDSTWVP 357
696 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
697 299 SITDIQINGIPTWGVFTQLNSIPGSVAVCAFSMDIEKVFKEQKTPDSVMTAVP 358
698 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
699 358 DERVPKPRGCCAGSSSLERYATSNFPDCTLNFIKTHPMDENVPSINRPFRLTMR 417
700 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
701 359 EDKVPKPRGCCAGKGLAEVATSDPDETLSEFKSHPLXDSAVPEIADPEWPKTEVP 418
702 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
703 418 YKTKAVDTNACPVNHWVFGSGEKGIIKELAPGNSGFINKSLFLEMSVYNSEK 476
704 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
705 419 YKTLAISVDHSAGPVQNYTVIFVSGEAGVVKVJAK--TSFPGKDSVLEEBEINYNK 476
706 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 477 CSYDGVEDKPIVGMQDRASSLVYAFSTCVIKVFLGRCERHCKKKTCTASHDFYQWI 536
708 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
709 477 CSAENSEDKKVLISLODKQHRLYVAFSSCIHRIPLGCEHYGSKKSCIASRDFYQWI 536
710 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
711 537 KEGGACSHLSPNSRLT-FEQDLERNTDGLGCHNSFVALNHSRSLSPSTTSSTAGH 595
712 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
713 537 SC-GSGCRPTFGMLAEGYQCTEFKTHLGGCHG----- 570
714 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
715 596 GYASRSGKLDWKLHLLDPSDIDPLJAVSSHNCKGKIFESYLKGDOLYVETLLAV 655
716 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
717 571 -----VWKE--VSGSGS----- 593
718 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 656 ILAFVXGAVFSGITVYVCGD-HRRKDVAVYVORKEKETHSRAGSMESVTKSGLFG 710
720 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
721 594 PFAFVVGAFIAGVAVYCYRDVFKVQRK--HKDAESAQSCDSSGSKAKGLNGLPDSVK 651
722 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
723 711 DTGSKDPKPEALITPLMNGKILATPNTANNK--KADQHLQDUTALPTPESTPLQOKRP 770
724 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
725 652 EYQCNDCSKLNSLLTSKELPPNGCTKSNVMDHRRGQPPLEALPTPESTPLVQLKTLQ 711
726 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
727 771 SRGSRWEMENQNLINACTKDYDPM3SPVLTDLPLRASPH-PSVTVLPITCCQYCHEY 829
728 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 712 AMKSHSEKAGH--GASRRETQFQFSSPPPHSPJ--SKGHIPSVIVPNATHYNTSFS 767

```

```

418 YRUTKIAVDTAGAPYORHYTYVFLSEKQITLKFAIRIGNSGF-LNDSLFLEKSNVFNSEK 476
419 YRUTALSDKRSAGFYQNYTVIFVSEAGNVLKFAK-TSPFSLNDSVLLEEIEAYNHAK 476
477 CSYDGVDFKRMKIMVCLDRASSLYVAFSTVIKVFLORCERHCKKTKIASECPYCGMI 536
477 CSAENEDKKVLSIQLOKHIALYVAFSSCTIRIPLSRCEYVGSCKSCIAERDPYCGMI 536
537 KEGGACSHLSPNSRLT-----FEQDIERGNTDGLGDCHNSFVALNGHSSSL 582
537 SQ-GSCGRVTPGXLLLTEDFPFAPNHHAEGVEQDTEFGNTAHJGDCRG----- 583
583 LPSTTTSDTAQEGYERSGGMLDKHLLDSDPDPJGAVSSHHNQKGV-RESYAKGH 642
584 -----VRME--VQSGES----- 593
643 DGLVEVTLALAVILAFMGAVFSGITVYVCD-HRKKVAVYQREKELTHSRNRSYSS 701
594 NQWVMVNLITCVFAMVVLGAFAGVAVYCYRDMFVKPKI--HKDAESAQSCITSSGS 651
702 VTKLSGLFG---DTQSKDKRKPPEAILTPMHNGKLTATPGNTAKMLIKADQHLHLLTALPT 757
652 FAKLNGLFDSEVKEVQNIQSPKYSKLLTSRAELPNSGDTKSNVMDHROQPELAALPT 711
758 PESTPTLQCKRKPGRSREMERONQNLINACTKQVPMGSPVITDPLRASPSPHPSVAV 817
712 PESTPVLHCKTLQAMKSHSEKAGH--GASRKETPOFFSPSPPHSP--SHGHIPSAIV 778
818 LP-TQCGYQHEV-----VDQP---KQSEVAQMAEDQAATLEYVTKIHEHLS 860
768 LPNATHDYNTSFSNSNAHKAEEKLQNI-CHPDTKSSSEKDEHRSVDSRNTL--NLLKLHLN 836
961 SKSPN-----HGVN-JVENJDSL---PPKVPQREASLGPFGASLSQTGLSKSLEM 906
926 DPNSKPAIMGDIQMAHQNLMDNGSGVSEVFNKPRREASQSPSTIPNSPTKEVQV 965
907 HHSSSY---GVDFKSYFTNSLTSSQATTLKRNTHSSNSHL-SANQSFQSD-NEPPA 962
986 PTFGVNMTSLERQGHKNSQ--HISAPKK-NLKGNGVLSLQFSNMVRGYNKFTPT 943
963 PQRVDS-QVHSSQPSGCAVTVSRQPS-NAVNSLT-----RSGIKRTPSLKPPVPKPFP 1016
944 GAKVDYVQ-----GTPSVHLQPS-SRQSSITSKCTLPETQLKTFPSLKPPVPKPFP 966
1027 APUSTSMKP 1025
397 VPQTPSVRP 1005

```

ESULT 15

```

9P249 D D Q9P249 PRELIMINARY; PR7; 1022 AA.
C C Q9P249;
T T C1-CCT-2000 (TremBirel, 15, Created)
T T 01-CCT-2001 (TremBirel, 18, Last sequence update)
T T 01-MAR-2003 (TremBirel, 23, Last annotation update)
T T Hypothetical protein K1AA1479 (P-fragment).
N N K1AA1479.
S S Homo sapiens (Human);
S S Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelesstori;
C C Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
X X NCBI_TaxID=9606;
N N [1].
P P SEQUENCE FROM N.A.
P P MEDLINE=20277482; PubMed=10819333;
P P Nagase T., Kikuno R., Ishikawa K., Hirose M., Chaya G.,
P P "Prediction of the coding sequences of unidentified human
P P genes.XVII. The complete sequences of 100 new cDNA clones from brain
P P which code for large proteins in vitro.";
L L DNA Res. 7:143-150(2000).
R R EMBL; AB040912; BAA96003.2;
R R InterPro; IPR003653; P1exin-like.

```

```

DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
WK Hypothetical protein.
PT NON TER.
SQ SEQUENCE 1022 AA; 114372 MW; BE4FBD5EA02C69C4 CRC64;

Query Match 42.5%; Score 2314; DB 4; Length 1022;
Best Loca. Similarity 45.4%; Pred. No. 3,5e-179;
Matches 494; Conservative 155; Mismatches 292; Indels 148; Gaps 26;

```

```

QY 1 MRSALLLYFTLL---HPAGAGFPEDSEPISSISGNTKQYPIVFGHKPGQNTTCRRHLD 57
DE 12 MRFVLLAYLILLMWSGLRAVSPFDEDEPLNTVDYHSRQYVFRG-RPSONESC-HRJD 69
CY 59 LOMIMXNGTLYIAARDHIVVDJDTSHTEIYCSKLLTWKSRQADVETGRMKKIKKDEC 117
DE 70 FQMLCKIRDTILYIAGRQOVTVNLNEMPKTEVLPKALTWRSQQQRENCAYKGGKHKDEC 129
QY 118 HNFIRVLKKNDADLFVCGTNAPFSPCRNYKQCTLEPFGDEFSGMARCPYDAKHANVALF 177
DE 130 HSF-KVFPFRNDENVFVCGTNAPFNEMCRVYVRLSTLEYDGEISGLARCPFDARQTNVALF 189
QY 178 ADGKLYSATVDFALDAVYNSLGSPTLPTWKHDSKWLKPEYFVCAVQYGYIYEFER 237
DE 190 ADGKLYSATVADFASDAVYRSMGCGSALRTIKYDSKWKEPHFLHAIEYGNFVYFFER 249
QY 238 EIAVEYNIMKGVFPRVAQVCKNDYGGSQVLEKQWTSFLKARLNCVSPDSSHFYFNILQ 297
DE 250 EIAVEHNHKGAVYSRIARCKNDYGGSQVLEKHWTSFLKARLNCVSPDSSHFYFDVLQ 309
QY 278 AVTVIRINGRDVYATPSTPYNSIPGSAVCAVMDLIASVPTGRFKEQKSPSTWTVPV 357
DE 310 SITDIQINGIPTVGVFTFQINSIPGSAVCAPSMODIEKVFKEQKTPGSAWTAVP 369
QY 358 DERVKKPRPGCCAGSSSELEVATSNBPDDTAPFKTHPLNDAVPSIFNRPKFLRMVR 417
DE 370 BQVKKPRPGCCAKHGJAEAYKTSIDPDELTFIKSHPLNCSAVPPAREPWTKTRVR 429
QY 418 YEDTKIADTAAGPYQKHTVWFLQSEKGIILKFLARIGNSGF-LNDSLFLEKSNVFNSEK 476
DE 430 YELNAISVDHAGAGYQNYTVIFVSEAGNVLKFAK-TSPFSLNDSVLLEEIEAYNHAK 487
QY 477 CSYDGVDFKRMKIMVCLDRASSLYVAFSTVIKVFLORCERHCKKTKIASECPYCGMI 536
DE 488 CSAENEDKKVLSIQLOKHIALYVAFSSCTIRIPLSRCEYVGSCKSCIAERDPYCGMI 547
QY 537 KEGGACSHLSPNSRLT-----FEQDIERGNTDGLGDCHNSFVALNGHSSSL 582
DE 548 SQ-GSCGRVTPGXLLLTEDFPFAPNHHAEGVEQDTEFGNTAHJGDCRG----- 594
QY 583 LPSTTTSDTAQEGYERSGGMLDKHLLDSDPDPJGAVSSHHNQKGV-RESYAKGH 642
DE 595 -----VRME--VQSGES----- 604
QY 643 DGLVEVTLALAVILAFMGAVFSGITVYVCD-HRKKVAVYQREKELTHSRNRSYSS 701
DE 655 NQWVMVNLITCVFAMVVLGAFAGVAVYCYRDMFVKPKI--HKDAESAQSCITSSGS 662
QY 702 VTKLSGLFG---DTQSKDKRKPPEAILTPMHNGKLTATPGNTAKMLIKADQHLHLLTALPT 757
DE 663 FAKLNGLFDSEVKEVQNIQSPKYSKLLTSRAELPNSGDTKSNVMDHROQPELAALPT 722
QY 758 PESTPTLQCKRKPGRSREMERONQNLINACTKQVPMGSPVITDPLRASPSPHPSVAV 817
DE 723 PESTPVLHCKTLQAMKSHSEKAGH--GASRKETPOFFSPSPPHSP--SHGHIPSAIV 778
QY 818 LP-TQCGYQHEV-----VDQP---KQSEVAQMAEDQAATLEYVTKIHEHLS 860
DE 779 LPNATHDYNTSFSNSNAHKAEEKLQNI-CHPDTKSSSEKDEHRSVDSRNTL--NLLKLHLN 836
QY 861 SKSPN-----HGVN-JVENJDSL---PPKVPQREASLGPFGASLSQTGLSKSLEM 906

```

837 DPNRPAWGOIQXHCNVLDFKXGSMSEVIFKVPKPRNSVSPSTLPRKSPTRVUW 896  
837 HESSY---GVYKRSYFTNTEFSHCATTCKENNTNSSSHLGRNOSFGEQ-NEEPA 962  
837 PTPGVEMTSLERQGYHKNSQF-HSISAVEK-NLKSPIGVLGRQPSMNEGUMPTPT 954  
837 PBYVDSIOCHSSQFSQAVTVSRQPSLNANSLT-----SSHLKETSLSKDEVPKRSF 1014  
837 GAKVGYIO-----GTPSVHLQPSLSPQSSYTSNGTLPTSLKTESLKEVPKRSF 1007  
1017 APJSTSMXP 1025  
1008 VPQTPSVRP 1016

Search completed: October 23, 2003, 17:12:38  
Job time : 65 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Computer: 116.

protein - protein search, using sw model

Run on: October 23, 2003, 17:09:37 ; Search time 10 Seconds  
(without alignment)  
3259.767 Million cell: updates: 900

```

title: US-09-856-681-2
rfect score: 5450
sequence: 1 MRSSAALLLYFTLLHENGAGE .....EKPQAPLSTINKNKDANT 1030

```

oring table: BLCUM62 Gapod 10.0 , Gapext 0.5

```

searched:      1:07663 seqs, 158726573 residues
data: number of hits satisfying chosen parameters: 1:87863

```

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                  Maximum Match 100
                  Plisted first 45

```

```

database :
A_Geneseq_19Jun03.*
1: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1980.DAT.*
2: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1981.DAT.*
3: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1982.DAT.*
4: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1983.DAT.*
5: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1984.DAT.*
6: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1985.DAT.*
7: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1986.DAT.*
8: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1987.DAT.*
9: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1988.DAT.*
10: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1989.DAT.*
11: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1990.DAT.*
12: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1991.DAT.*
13: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1992.DAT.*
14: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1993.DAT.*
15: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1994.DAT.*
16: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1995.DAT.*
17: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1996.DAT.*
18: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1997.DAT.*
19: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1998.DAT.*
20: /SID1/cgcdata/geneseq/Genesecp-emb1/AA1999.DAT.*
21: /SID1/cgcdata/geneseq/Genesecp-emb1/AA2000.DAT.*
22: /SID1/cgcdata/geneseq/Genesecp-emb1/AA2001.DAT.*
23: /SID1/cgcdata/geneseq/Genesecp-emb1/AA2002.DAT.*
24: /SID1/cgcdata/geneseq/Genesecp-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SULT No.	Query Match	Score	Length	DB	ID	Description
1	5450	100.0	1030	21	AAY71460	Human semaphorin 6
2	5103.5	93.6	975	22	AAB590731	Human CJI45.1 pro
3	5088	93.4	974	19	AAW64221	Human secreted pro
4	4921	90.3	939	21	AAB23030	Human semaphorin p
5	4590.5	84.2	884	21	AAE23031	Human semaphorin p
6	3677	67.5	699	22	AAB55139	Human protein sequ
7	3373	61.9	630	21	AAB23043	Human semaphorin p
8	2980.5	54.7	574	22	AAB55317	Human protein sequ
9	2952	54.2	562	22	AAB54104	Human protein sequ



e.g. cytokine and cell proliferation/differentiation activity, the immune system and hematopoiesis regulating activity.

Disclosure: Page 487-490; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted proteins AAF96667 - AAF98075. The cDNA clones are isolated from various tissue types, and may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate protein expression. The polypeptides and nucleic acids may be used as nutrients or to modulate cytokine and cell proliferation/differentiation activity and may also be involved in regulation of the immune system. The cDNA sequences, proteins, their agonists and/or antagonists exhibit hematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity; receptor/ligand activity; anti-inflammatory activity; hematopoiesis activity; cachectin/tumour suppressor activity; and/or tumour inhibition activity. Included in the invention are probes represented in AAF98490 - AAF98572 which are specific for the cDNA clones encoding the secreted proteins.

Sequence 975 AA:

Query Match 93.6%; Score 5103.5; DB 22; Length 975;

Best Local Similarity 94.4%; Pred. No. 2;

Matches 972; Conservative 0; Mismatches 3; Indels 55; Gaps 1;

```

1 MSSEALLLYFTLLHFAGAGPEDESEPTISHSNYTKQYVFVGHKPGKNTTORHRLDQ 60
1 MSSEALLLYFTLLHFAGAGPEDESEPTISHSNYTKQYVFVGHKPGKNTTORHRLDQ 60
61 IMINNGTLYIARDHIYTVLDTISHTTEIYCSKLTWKSROAEVDTORVKGKHDCSNF 120
61 IMINNGTLYIARDHIYTVLDTISHTTEIYCSKLTWKSROAEVDTORVKGKHDCSNF 120
121 IKVLLKKNDJALFVGTNAFNPCSNYKMDTLEPGDEFSGMARCPYDAKANYALPADG 180
121 IKVLLKKNDJALFVGTNAFNPCSNYKMDTLEPGDEFSGMARCPYDAKANYALPADG 180
181 KLYSATVTDFAIDAVIRSLGSEPTRTVTHCSKWLKEPVEVQAVYGVGYIFFFREIA 240
181 KLYSATVTDFAIDAVIRSLGSEPTRTVTHCSKWLKEPVEVQAVYGVGYIFFFREIA 240
241 VEYNTWGVKVPFPRVACVKNVGGGQVLEKQWTSFLKAPLNCSPGGSHFVNTLOAVT 300
241 VEYNTWGVKVPFPRVACVKNVGGGQVLEKQWTSFLKAPLNCSPGGSHFVNTLOAVT 300
301 DVIRINGROWLATFTSTPYN:PGSAVCAVYDMLD:ASVFTGSEKQCFSPDSTWTPVDEP 360
301 DVIRINGROWLATFTSTPYN:PGSAVCAVYDMLD:ASVFTGSEKQCFSPDSTWTPVDEP 360
361 VPKPRPGCCAGSSLEPYATSNFEDDTNFKTPLMDEAVPS:FNENFELRTWYVRL 420
361 VPKPRPGCCAGSSLEPYATSNFEDDTNFKTPLMDEAVPS:FNENFELRTWYVRL 420
421 TKIADVTAGAPQNHVTVFVJGSEKGIKFKFARI:GNSGFNDLSLPLEMSVYVSEKQCYD 480
421 TKIADVTAGAPQNHVTVFVJGSEKGIKFKFARI:GNSGFNDLSLPLEMSVYVSEKQCYD 480
481 GVECKRINQXQDRASSLYVAFSTCV:KVP:GRGERHKGKXKTKIARDPYGVNMYEKG 540
481 GVECKRINQXQDRASSLYVAFSTCV:KVP:GRGERHKGKXKTKIARDPYGVNMYEKG 540
541 ACSHLSPNRLTFEOD:ERGVNTEGLGOCNHSFVALNCHSSSLDPTTSCSTAGQGVSR 600
541 ACSHLSPNRLTFEOD:ERGVNTEGLGOCNHSFVALNCHSSSLDPTTSCSTAGQGVSR 600
601 GGVLDWKHLLDSDPTDPLGAVSSHQKKGKVIRESY:KGHDQLVPVTLIAVILAFV 660
601 GGVLDWKHLLDSDPTDPLGAVSSHQKKGKVIRESY:KGHDQLVPVTLIAVILAFV 660
661 NGAVFSGITVYVCDHRRKDVAVVQRKEKELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 720
661 NGAVFSGITVYVCDHRRKDVAVVQRKEKELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 720

```

```

Db      606 MGAVRPSGITVYVCDHRRKDVAVVQRKEKELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 665
Qy      721 AILTPXNHGKATPONTAKMLKADQHLLDTALTPTPESTPTLQOKRXPDSRSEWERN 780
Db      666 AILTPXNHGKATPONTAKMLKADQHLLDTALTPTPESTPTLQOKRXPDSRSEWERN 725
Qy      781 QNLINACTKMPFMGSPV:PTDLPLRASPSH:PSVYVLPITQGYCHEYVDCPKSEVAC 840
Db      726 QNLINACTKMPFMGSPV:PTDLPLRASPSH:PSVYVLPITQGYCHEYVDCPKSEVAC 785
Qy      841 MALEDOQAATLEYNKTIKHEH:SSKSPNKGVMJVENLDSLPFKVPQREASLGPQGCASLQ 900
Db      786 MALEDOQAATLEYNKTIKHEH:SSKSPNKGVMJVENLDSLPFKVPQREASLGPQGCASLQ 845
Qy      901 SXPLEMHHSYGVQVYKRSYPTNS:TRSHOATLKNNTNSNSHLSNQSFGGDNTP 960
Db      846 SXPLEMHHSYGVQVYKRSYPTNS:TRSHOATLKNNTNSNSHLSNQSFGGDNTP 905
Qy      961 PAPCRVDSIQVHSSQPSGCAVTVSRQPS:KAYNSLTPSGLKXETPSLKPDVPPKSFAPLS 1020
Db      906 PAPCRVDSIQVHSSQPSGCAVTVSRQPS:KAYNSLTPSGLKXETPSLKPDVPPKSFAPLS 965
Qy      1021 TSMKPNDACT 1030
Db      966 TSMKPNDACT 975

RESULT ?
AAK6422:
ID AAK6422: standard; Protein: 974 AA.
XX
AC AAK6422:
XX
DT 06-OCT-1998 (first entry)
XX
DE Human secreted protein from clone C0145_1.
XX
KW Secreted protein; human fetal brain; nutrition; cytokine; stimulant;
KW cell proliferation; differentiation; immune system; suppressor; ligand;
KW regulator; hematopoiesis; tissue growth; activin; inhibin; haemostatic;
KW chemotaxis; chemokinetic; thrombosis; receptor; cadherin; tumour;
KW anti-inflammatory.
XX
OS Homo sapiens.
XX
DX WD0987205-A2.
XX
PD 25-JUN-1998.
XX
PF 17-DEC-1997; 97MO-US23330.
XX
PR 16-DEC-1997; 97US-0991872.
PR 18-DEC-1996; 96US-0769192.
PR 13-JAN-1997; 97US-0783401.
XX
PA GENEV : GENETICS INST INC.
XX
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JX, Merberg D;
PI Racie SA, Spaulding W, Treacy N;
XX
DR WP: 1998-362774/31.
DR N-PSDS; AAV44295.
XX
PT New polynucleotides and secreted proteins - obtained from human
PT foetal brain, human adult testes, human adult brain and human adult
PT salivary gland cDNA libraries
XX
PS Claim 173; page 71-74; l10pp; English.
XX
CC This sequence represents a novel secreted protein from clone C0145_1
CC isolated from a human fetal brain cDNA library. This protein has
CC applications for nutritional use, cytokine and cell
CC proliferation/differentiation activity, immune stimulating or

```

C suppressing activity, hematopoiesis regulating activity, tissue growth  
C activity, activin/inhibin activity, chemotactic/chemokinetic activity,  
C haemostatic and thrombotic activity, receptor/ligand activity,  
C anti-inflammatory activity, cadherin/tumour invasion suppressor activity,  
C tumour inhibition activity and other activities.

X Query Match 93.4%; Score 5089; DB 19; Length 974;  
X Best Local Similarity 94.3%; Pred. No. 2;  
X Matches 971; Conservative 0; Mismatches 3; Indels 56; Gaps 2;

1 MRSEALLLYFTLLHFAAGPEDESEPTISNGNYTKQYVVFVGHKPGRNTPQHRJLQIM 60  
b 1 MRSEALLLYFTLLHFAAGPEDESEPTISNGNYTKQYVVFVGHKPGRNTPQHRJLQIM 60  
/ 61 INIKNGTLYAARDHIYTVVIDTSHTSEIYCSKLTWKSSQADVDTCRMKGKHKECHNP 120  
b 61 INIKNGTLYAARDHIYTVVIDTSHTSEIYCSKLTWKSSQADVDTCRMKGKHKECHNP 119  
/ 121 IKVLLKXNDALFYCGTNAFNSCRNTKQJDTLFPQDEFSGMARCPYDAHANVALFADG 180  
b 121 IKVLLKXNDALFYCGTNAFNSCRNTKQJDTLFPQDEFSGMARCPYDAHANVALFADG 179  
/ 181 KLYSATVTDPLAIDAVYVSGSGSEPTRTYKHCSKCLKEPYVCATYCYGYVFFPREIA 240  
b 181 KLYSATVTDPLAIDAVYVSGSGSEPTRTYKHCSKCLKEPYVCATYCYGYVFFPREIA 239  
/ 241 VEYNTMGKVFVFFVAQCKKMSQSVJESKQNTSFLKAPLNSGVFGDGHFPNLIQAVT 300  
b 241 VEYNTMGKVFVFFVAQCKKMSQSVJESKQNTSFLKAPLNSGVFGDGHFPNLIQAVT 299  
/ 301 DVIRINGDQVLTFTSPYNSIPGSAVCATQMDLQASVFTQREKESKQSDSTWTVPEER 360  
b 301 DVIRINGDQVLTFTSPYNSIPGSAVCATQMDLQASVFTQREKESKQSDSTWTVPEER 359  
/ 361 VPKRPPGCCAGSSLEAYATSNBPFDTLNFITKPLMDCAVPSIFNRMWELRTVVEYEL 420  
b 361 VPKRPPGCCAGSSLEAYATSNBPFDTLNFITKPLMDCAVPSIFNRMWELRTVVEYEL 419  
/ 421 TKIADVTNAGYQVHTVYVFGSNGIILKELARIGNSGFLNDSFILEXSVYNSKQSDY 480  
b 421 TKIADVTNAGYQVHTVYVFGSNGIILKELARIGNSGFLNDSFILEXSVYNSKQSDY 479  
/ 481 GVEKRIKMGVGLDASSLLVAFSTVYKPLGRGSRHGMKCKXTQASRQPCGMWKEG 540  
b 481 GVEKRIKMGVGLDASSLLVAFSTVYKPLGRGSRHGMKCKXTQASRQPCGMWKEG 539  
/ 541 ACSHLSFNSRLTPFDIERNTDCLGSCNHSFVALACHSSSELPSTTTSDSKAQGEYER 600  
b 541 ACSHLSFNSRLTPFDIERNTDCLGSCNHSFVALACHSSSELPSTTTSDSKAQGEYER 599  
/ 576 -----GVRESYLKGHQLPVFLALAVLAV 604  
b 576 -----GVRESYLKGHQLPVFLALAVLAV 604  
/ 651 MGAVFSGITVYVCVCHRRKCVAVYVORKEKSLTHSRGSGMSVTKLGLFQDTOSKDPKE 720  
b 651 MGAVFSGITVYVCVCHRRKCVAVYVORKEKSLTHSRGSGMSVTKLGLFQDTOSKDPKE 719  
/ 721 AILPLMNGKLTATPGNTAKMLKACHHLDLIALTPPESTPTIQQRKXPSGSRWERN 780  
b 721 AILPLMNGKLTATPGNTAKMLKACHHLDLIALTPPESTPTIQQRKXPSGSRWERN 779  
/ 781 QNLINACTKOMPXGSPVPTDGLFASPSHISVWVLPITQOQYCHEYVDOPKSEVAQ 840  
b 781 QNLINACTKOMPXGSPVPTDGLFASPSHISVWVLPITQOQYCHEYVDOPKSEVAQ 839  
/ 841 MAJEDQATHEYKTIKEHLSKSNFHNGLVENLGLSPTPKVPQREASLAPPASLSQGL 900  
b 841 MAJEDQATHEYKTIKEHLSKSNFHNGLVENLGLSPTPKVPQREASLAPPASLSQGL 899  
/ 901 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKNNNTNSNSHLSRNSQSGRGDNP 960

845 SKLEMHHSFSGVGVYKRSYPTNSLTRSHQATTLKNNNTNSNSHLSRNSQSGRGDNP 904  
961 PAPQRVDSIQVHSSQSGGAVTVSRQPSLNAYNSLTRSGLRKTPSLKDPVFKPSPFAPLS 1020  
905 PAPQRVDSIQVHSSQSGGAVTVSRQPSLNAYNSLTRSGLRKTPSLKDPVFKPSPFAPLS 964  
1021 TSMKPNDACT 1030  
965 TSMKPNDACT 974

RESULT 4  
AAB2303C  
ID AAB2303C standard; Protein: 939 AA.  
XX AAB2303C;  
XX 16-JAN-2001 (first entry)  
XX Human seraphorin protein-like splice variant, SECX 2864913-1.  
XX SECX protein; human; secreted; membrane-associated; cancer;  
XX proliferation regulator; differentiation regulator; non-malignant tumour;  
XX immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
XX infection; inflammatory disorder; arthritis; haematopoietic disorder;  
XX skin disorder; cardiovascular disorder; atherosclerosis; stenosis;  
XX neurological disease; Alzheimer's disease; trauma; wound;  
XX spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
XX anti-HIV; anti-inflammatory; anti-arthritis; antiarteriosclerotic;  
XX neuroprotective; vulnery; anti-allergy; antimicrobial; cardiac;  
XX dermatological; gene therapy.  
XX Homo sapiens.  
XX WC0200053742-A2.  
XX 14-SEP-2000.  
XX 09-MAR-2000; 2000WJ-USC6280.  
XX 09-MAR-1999; 99US-0123667.  
XX 06-MAR-2000; 2000US-0123667.  
XX (CJUA-1) CURAGEN CORP.  
XX Shimkats RA;  
XX MPI; 2000-5943:B/56.  
XX N-PSDB; AAA93617.  
XX Novel human membrane associated or secreted polypeptides and  
XX polynucleotides useful for diagnosis, prevention and treatment of  
XX pathological states such as cancer, immune, cardiovascular and  
XX neurological disorders.  
XX Claim 1; Fig 2; 151pp; English.  
XX Sequences AAB23029-521049 represent human SECX proteins. The SECX  
XX proteins of the invention are either secreted or membrane-associated  
XX proteins and act as regulator of cellular proliferation and  
XX differentiation. SECX proteins or nucleotides are useful for diagnosing  
XX the presence of, or predisposition to, a disease associated with altered  
XX levels of SECX proteins and nucleotides. The SECX proteins are also  
XX useful to screen compounds that modulate SECX activity or expression. The  
XX interaction of a SECX protein with other cellular proteins may be useful  
XX to modulate the activity of a partner protein, cellular proliferation,  
XX cellular differentiation and cell survival. SECX nucleotides are useful  
XX for the recombinant expression of SECX protein, and may be used to detect  
XX SECX mRNA or genetic lesions in the SECX gene. They may also be used to  
XX modulate SECX expression (e.g., using antisense oligonucleotides). SECX  
XX nucleic acid sequences are also useful for identifying a cell or tissue  
XX type in a biological sample, and in forensic biology. SECX primers or



probes are useful for detecting the presence of SECC nucleotides and for screening tissue cultures for contamination. Diseases that may be treated or prevented using SECC proteins or nucleotides include cancer (e.g., colorectal carcinoma, prostate cancer), benign tumors, immune disorders (including autoimmune diseases, transplant rejection, allergies, AIDS), infections, inflammatory disorders, arthritis, haematopoietic disorders, skin disorders, cardiovascular disorders, atherosclerosis, restenosis, neurological diseases (e.g., Alzheimer's disease), trauma (e.g., surgical or traumatic wounds, spinal cord injury), and skeletal disorders.

Q	Query Match	90.3%	Score	4921	DR	2	Length	933
Q	Best Local Similarity	99.5%	Pred	No. 0				
Q	Matches	928	Conservative	2	Mismatches	2	Indels	0
Q	Gaps	0						
Q	Sequence	933 AA:						
Q	1	MRSEALLVFTLLHFGAGAGFPDSEPIGISHGNYTKQIVFVGHKPGGNTTQRHRLDTCM	60					
Q	1	MRSEALLVFTLLHFGAGAGFPDSEPIGISHGNYTKQIVFVGHKPGGNTTQRHRLDTCM	60					
Q	61	IMNMGTLVIAARDHIYVDIOTSHTESYGSKLTWKSRAQADVTCTMKYKHDECHNF	120					
Q	61	IMNMGTLVIAARDHIYVDIOTSHTESYGSKLTWKSRAQADVTCTMKYKHDECHNF	120					
Q	121	IKVLKXKDDALFVCGTNAFNSCENYKQOTLPPFGDSEFGMARCPYDAKXANVALPADG	180					
Q	121	IKVLKXKDDALFVCGTNAFNSCENYKQOTLPPFGDSEFGMARCPYDAKXANVALPADG	180					
Q	181	KLYSATVTDPLAIDAVIYRSLGESPTLTKVHDSKMLKEPYFVQAVYGVYIYFFFRRIA	240					
Q	181	KLYSATVTDPLAIDAVIYRSLGESPTLTKVHDSKMLKEPYFVQAVYGVYIYFFFRRIA	240					
Q	241	VEYNTMGKVYFPRVAQVKNQDGGGQVLEKQNTSFLKAPLNCSPVGGSHFYFTLQVAT	300					
Q	241	VEYNTMGKVYFPRVAQVKNQDGGGQVLEKQNTSFLKAPLNCSPVGGSHFYFTLQVAT	300					
Q	301	DVIRINGRDVLATSTPVNSTPGSAVCAVMDLTSVETGPFKECKSPDSTWTPVDPDR	360					
Q	301	DVIRINGRDVLATSTPVNSTPGSAVCAVMDLTSVETGPFKECKSPDSTWTPVDPDR	360					
Q	361	VKPRPGCCAGSSSLERYATSNFEEDDTNFKTHFLMDEAVPSFNEPFWELRWVYVRL	420					
Q	361	VKPRPGCCAGSSSLERYATSNFEEDDTNFKTHFLMDEAVPSFNEPFWELRWVYVRL	420					
Q	421	TKIAYDTAAGFYQHTYVPLGSEKGIILKFLARINSGFUNDLSLEKSVYNSKEKCYC	480					
Q	421	TKIAYDTAAGFYQHTYVPLGSEKGIILKFLARINSGFUNDLSLEKSVYNSKEKCYC	480					
Q	481	GVEDKFINSGDRASSSLYVAFSTCVTKVHLGRERKQKPKTCIAGRDYGVKWIKEGG	540					
Q	481	GVEDKFINSGDRASSSLYVAFSTCVTKVHLGRERKQKPKTCIAGRDYGVKWIKEGG	540					
Q	541	ACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTTSDTAEGVSEK	600					
Q	541	ACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTTSDTAEGVSEK	600					
Q	601	GXLPLKWLJDSPOSTDLGAVSSNNHCKKQVRESYKXGHDQVVPVTLIAVIAVAFV	660					
Q	601	GXLPLKWLJDSPOSTDLGAVSSNNHCKKQVRESYKXGHDQVVPVTLIAVIAVAFV	660					
Q	661	MGAVSGITVYVCGDHRKRVAVVCKKELNLSRERGYSVVKLSGFGTGQKRPKPE	720					
Q	661	MGAVSGITVYVCGDHRKRVAVVCKKELNLSRERGYSVVKLSGFGTGQKRPKPE	720					
Q	721	AILTLPJMXNGXKATGNTAKMLIKADQCHLDELTAJPTPESTTLQCKRPKPGSGREWEEN	780					
Q	721	AILTLPJMXNGXKATGNTAKMLIKADQCHLDELTAJPTPESTTLQCKRPKPGSGREWEEN	780					
Q	781	QNLINACTKDMPPKSGPVIFTDJFLRAGSPSHIPSVYVLPITCGGQVQHYVYVQPKMSEVAG	840					
Q	781	QNLINACTKDMPPKSGPVIFTDJFLRAGSPSHIPSVYVLPITCGGQVQHYVYVQPKMSEVAG	840					

841 MALEDQAATLEKYTIKEHLSSKSPNGVNLVENDLSLPKVPQREASLGPASLSQTGL 900  
 841 MALEDQAATLEKYTIKEHLSSKSPNGVNLVENDLSLPKVPQREASLGPASLSQTGL 900  
 901 SKLELXHSSSSGVGYKRSYPTNLSLRSHQAT 932  
 901 SKLELXHSSSSGVGYKRSYPTNLSLRSHQAT 932

RESULT 5  
 AAB23031  
 ID AAB23031 standard; Protein: 694 AA.  
 AC AAB23031:  
 DT 16 JAN-2002 (first entry)  
 DE Human semaphorin protein-like splice variant, SECC 2864933-2.  
 KW SECC protein; human; secreted; membrane-associated; cancer;  
 KW proliferation regulator; differentiation regulator; non-malignant tumor;  
 KW immune disorder; autoimmune disease; transplant rejection; AIDS;  
 KW infection; inflammatory disorder; arthritis; haematopoietic disorder;  
 KW skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
 KW neurological disease; Alzheimer's disease; trauma; wounding;  
 KW spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
 KW anti-HIV; antiinflammatory; antithrombotic; antitumorsclerotic;  
 KW neuroprotective; valvulopathy; antiatherogenic; antimicrobial; cardiac;  
 KW dermatological; gene therapy.

XX Homo sapiens.  
 OS  
 XX W0200051742-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF C9-MAR-2000; 2000MO-US06260.  
 XX  
 PR C9-MAR-1999; 99US-0123667.  
 PR C9-MAR-2000; 2000US-0123667.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Simkrets RA;  
 XX  
 XX WPI; 2000-594318/56.  
 DR N-PSDB: AAA230618.  
 XX  
 PT Novel human membrane associated or secreted polypeptides and  
 PT polynucleotides useful for diagnosis, prevention and treatment of  
 PT pathological states such as cancer, immune, cardiovascular and  
 PT neurological disorders.

Claim 1: Fig 3; 15:pp; English.

Sequences AAB23029-B23048 represent human SECC proteins. The SECC proteins of the invention are either secreted or membrane-associated proteins and act as a regulator of cellular proliferation and differentiation. SECC proteins or nucleotides are useful for diagnosing the presence of, or predisposition to, a disease associated with altered levels of SECC proteins and nucleotides. The SECC proteins are also useful to screen compounds that modulate SECC activity or expression. The interaction of a SECC protein with other cellular proteins may be useful to modulate the activity of a partner protein, cellular proliferation, cellular differentiation and cell survival. SECC nucleotides are useful for the recombinant expression of SECC protein, and may be used to detect SECC mRNA or genetic lesions in the SECC gene. They may also be used to modulate SECC expression (e.g., using antisense oligonucleotides). SECC nucleic acid sequences are also useful for identifying a cell or tissue type in a biological sample, and in forensic biology. SECC primers or probes are useful for detecting the presence of SECC nucleotides and for screening tissue cultures for contamination. Diseases that may be treated or prevented using SECC proteins or nucleotides include cancer (e.g.,

C Colorectal carcinoma, prostate cancer, benign tumors, immune disorders  
C (including autoimmune diseases, transplant rejection, allergies, AIDS),  
C infections, inflammatory disorders, arthritis, rheumatoid arthritis,  
C skin disorders, cardiovascular disorders, atherosclerosis, restenosis,  
C neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,  
C surgical or traumatic wounds, spinal cord injury), and skeletal  
C disorders.  
X  
Q Sequence 894 AA:  
Query Match 84.2%; Score 4590.5; DB 2; Length 894;  
Best Local Similarity 93.7%; Pred. No. 0;  
Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;  
Y 1 MEEALLLYPTLAFAGAGPEDEPFISISGHGNTVTCQYVFGKFKGRNTTQHRLLDGM 60  
E 1 YRSALJJYFTLLAFAGAGPEDEPFIS:SGHNTVTCQYVFGKFKGRNTTQHRLLDGM 60  
Y 61 IMINGTJYNAARDHITVDITSHTEETVCSKLLTWKSPQAVDTCTCKKGRHDECHNF 120  
D 61 IMINGTJYNAARDHITVDITSHTEETVCSKLLTWKSPQAVDTCTCKKGRHDECHNF 120  
Y 121 IKVILKKKDALFYCGNAPNESLNNKVTLEFFGDEFSGMARCTYDAKHANVALPADG 180  
D 121 IKVILKKKDALFYCGNAPNESLNNKVTLEFFGDEFSGMARCTYDAKHANVALPADG 180  
Y 181 KYSATVITDLDAVLYRSLGESPTLTVKHDSKMLKEPYFVQAVDYGYEYFFPREIA 240  
b 181 KYSATVITDLDAVLYRSLGESPTLTVKHDSKMLKEPYFVQAVDYGYEYFFPREIA 240  
Y 241 VEYTMKATVFPVACVCKNDGSGORVLEKQNTSELKARLNCVPGDSHFVENILQAVT 300  
D 241 VEYTMKATVFPVACVCKNDGSGORVLEKQNTSELKARLNCVPGDSHFVENILQAVT 300  
Y 301 DVIRINDVYLATFTFYNS:PGSAYCAYMLASVFTGTFKEKQKQDSMTVPDESR 360  
D 301 DVIRINDVYLATFTFYNS:PGSAYCAYMLASVFTGTFKEKQKQDSMTVPDESR 360  
Y 361 VPKRPGCCAGSSLEAYATSNFPPDILFIKTHPLNDHANESIENRPFJBTMRYSL 420  
D 361 VPKRPGCCAGSSLEAYATSNFPPDILFIKTHPLNDHANESIENRPFJBTMRYSL 420  
Y 421 TKIADVTANGYQNTVYVGSSEKGIILKELARLGNCHLNDLFLFEMSVNSEXCSVD 480  
D 421 TKIADVTANGYQNTVYVGSSEKGIILKELARLGNCHLNDLFLFEMSVNSEXCSVD 480  
Y 481 QVEDKRVGMGLDRASSSYVAFSTCVIKVPLGCEPHGKCKTCTIAROPYCGMKEGG 540  
D 481 QVEDKRVGMGLDRASSSYVAFSTCVIKVPLGCEPHGKCKTCTIAROPYCGMKEGG 540  
Y 541 ACSNLSPNSRLTFEQDIEGNTDGG:GCHNSFALNCHASSLPSTTSDSTACQDYER 600  
D 541 ACSNLSPNSRLTFEQDIEGNTDGG:GCHNSFALNCHASSLPSTTSDSTACQDYER 600  
Y 601 GGMGLDWKLLSPDSTPLGAVSSHHQKNGVIRESYLKGHPQAVFVLALAVLAPV 660  
D 601 GGMGLDWKLLSPDSTPLGAVSSHHQKNGVIRESYLKGHPQAVFVLALAVLAPV 660  
Y 661 MGAVFSGITVYVCDHRKDVAVTCRKEKELTSPRSGMSVTKLSGLFECTOSKPKPE 720  
D 661 MGAVFSGITVYVCDHRKDVAVTCRKEKELTSPRSGMSVTKLSGLFECTOSKPKPE 720  
Y 721 ALLTPDMHNGKLATPGNTAKMLIADQHHLDLTALPTPESTFTLQQRKPSHSGREWRN 780  
D 721 ALLTPDMHNGKLATPGNTAKMLIADQHHLDLTALPTPESTFTLQQRKPSHSGREWRN 780  
Y 781 QNLINACTKNMPKGPSPVPTDGLFRASPSHPVWVLPITQQCYCHYYDQPKYSEVAC 840  
D 781 QNLINACTKNMPKGPSPVPTDGLFRASPSHPVWVLPITQQCYCHYYDQPKYSEVAC 840  
Y 841 MALEDCAATLEYKTIKEHLSSKSNHGNLVEN:DSLPPKVPQBEASLGPQASLSQTL 900  
D 841 MALEDCAATLEYKTIKEHLSSKSNHGNLVEN:DSLPPKVPQBEASLGPQASLSQTL 900

Qy 901 SKLEMHSSSYGVYKSPYPTNSLTRSHOAT 932  
Db 846 SKLEMHSSSYGVYKSPYPTNSLTRSHLT 877

RESULT 6  
AA095139  
ID AA095139 standard; Protein: 699 AA.  
XX AA095139:  
AC AA095139:  
XX 26-JUN-2001 (first entry)  
XX Human protein sequence SEQ ID NO:17154.  
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX Homo sapiens.  
XX EP:074617-A2.  
XX 07-FEB-2001.  
XX 28-JUL-2000; 2000EP-0116:26.  
XX 29-JUL-1999; 99JP-0248026.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0193767.  
PR 09-JUN-2000; 2000JP-0241895.  
XX (HELI-) HELIX RES INST.  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
X Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs -  
Claim 3: SEQ ID 17154; 2537pp + CD ROM; English.  
The present invention describes primer sets for synthesizing 5602  
full-length cDNAs defined in the specification. Where a primer set  
comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence/3'-end sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and  
in gene therapy. The primers are useful for synthesizing polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow obtaining of the full-length  
cDNAs easily without any specialised methods. AA03166 to AAH13628 and  
AAH13631 to AAH18742 represent human cDNA sequences; AA092446 to  
AA095893 represent human amino acid sequences; and AAH13629 to AAH13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.

Sequence 699 AA:  
Query Match 57.5%; Score 3677; DB 2; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1.6e-310;  
Query Match 57.5%; Score 3677; DB 2; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1.6e-310;

Matches 699, Conservative 0, Mismatches 0, Indels 0, Gaps 0

332 MLDIASVTGKEKECKSPDSTWTPDQBPVPRPGCCACAGSSLSRYATSNEPDTLNF 391  
|||||  
1 MLDIASVTGKEKECKSPDSTWTPDQBPVPRPGCCACAGSSLSRYATSNEPDTLNF 60  
|||||  
392 IKTHPLMDAVSIFENRWFUATVYVYETKIATVTAAGPYQMTWFGSEKSIILKFL 451  
|||||  
61 IKTHPLMDAVSIFENRWFUATVYVYETKIATVTAAGPYQMTWFGSEKSIILKFL 120  
|||||  
452 ARIKSGPLNDSLFLEKSVYNSKSCSDGVDKRMQMDRAUSSLYVAFSTCVIVP 511  
|||||  
121 ARIKSGPLNDSLFLEKSVYNSKSCSDGVDKRMQMDRAUSSLYVAFSTCVIVP 120  
|||||  
512 LGRCERHKGCKTCTASRPYCGWKKEGAGSHLSUNSLTFEQGIESQNTDGLGCENS 571  
|||||  
181 LGRCERHKGCKTCTASRPYCGWKKEGAGSHLSUNSLTFEQGIESQNTDGLGCENS 240  
|||||  
572 FYALAGHSSLSLFTTTSSTAQEGYERGGMLDKMILLDPESLTFEGAYSSNHQDXX 631  
|||||  
241 FYALAGHSSLSLFTTTSSTAQEGYERGGMLDKMILLDPESLTFEGAYSSNHQDXX 300  
|||||  
632 GVIRESYLKHQDLPVPTLAIATVLAIPWGAVFSGITVYVCDHRRKQDAVYORKKEL 691  
|||||  
301 GVIRESYLKHQDLPVPTLAIATVLAIPWGAVFSGITVYVCDHRRKQDAVYORKKEL 360  
|||||  
692 THSRRGSSSVTKLSLFGTOSKQPKPEALITPLVHNGKATPONTAKMLIKADQHLE 751  
|||||  
361 THSRRGSSSVTKLSLFGTOSKQPKPEALITPLVHNGKATPONTAKMLIKADQHLE 420  
|||||  
752 LTAJLTPTSTPTLOKRPKSGSRERNQNLINACTKQMPKMGSPVITPLIRASPSH 811  
|||||  
421 LTAJLTPTSTPTLOKRPKSGSRERNQNLINACTKQMPKMGSPVITPLIRASPSH 480  
|||||  
812 IPSVVLPITQGGYCHEYVQDPKSEVAQMALEDQAATLEYKIKHEHSSKSPHGNVLY 871  
|||||  
481 IPSVVLPITQGGYCHEYVQDPKSEVAQMALEDQAATLEYKIKHEHSSKSPHGNVLY 540  
|||||  
872 ENLDSPPKVPQREASLQPGASLSQTLGSKLEMHSSYGVYKSPYNTSLTRSHOA 931  
|||||  
541 ENLDSPPKVPQREASLQPGASLSQTLGSKLEMHSSYGVYKSPYNTSLTRSHOA 600  
|||||  
932 TTKRNNTNNSKSSLSRNCQSGRGNPPAPQCVDSIQVHSSQSGQAVTVSGQFSLNA 991  
|||||  
601 TTKRNNTNNSKSSLSRNCQSGRGNPPAPQCVDSIQVHSSQSGQAVTVSGQFSLNA 660  
|||||  
992 YNSLTRSGLKPTPSLKPDVPPKPSFAPLSTSMKPRDACT 1030  
|||||  
661 YNSLTRSGLKPTPSLKPDVPPKPSFAPLSTSMKPRDACT 699  
|||||

RESULT 7

AAB23043 standard; Protein; 630 AA.

X X AAB23043; Query Match: 61.9%; Score 3373; DP 21; Length 630;  
X X AAB23043; Best Local Similarity 99.8%; Pred. No. 3.8e-284;  
X X AAB23043; Matches 628; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

16-JAN-2001 (first entry)

Human semaphorin protein-like splice variant, SECX p02.1-2864933.

X X SECX protein; human; secreted; membrane-associated; cancer;  
X X proliferation regulator; differentiation regulator; not malignant tumour;  
X X immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
X X infection; inflammatory disorder; arthritis; haematopoietic disorder;  
X X skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
X X neurological disease; Alzheimer's disease; atherosclerosis; restenosis;  
X X spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
X X anti-HIV; anti-inflammatory; anaphylaxis; antineoplastic; cardiac;  
X X neuroprotective; valvular; anti-atherosclerotic;  
X X dermatological; gene therapy.

X X Homo sapiens.

XX PN WC0200053742-A2.  
XX XX 14-SEP-2000.  
XX XX 09-XAR-2000; 2000WC-0506280.  
XX XX 09-XAR-1999; 99US-0123667.  
XX XX 08-XAR-2000; 2000US-0123667.  
XX XX (CURA-) CURAGEN CORP.  
XX XX Shinketsu RA.  
XX XX WP1; 2000-554315/56.  
XX XX N-PSDB; AAA93630.  
XX XX Novel human membrane associated or secreted polypeptides and  
XX XX polynucleotides useful for diagnosis, prevention and treatment of  
XX XX pathological states such as cancer, immune, cardiovascular and  
XX XX neurological disorders.

XX XX Claim 1; Fig 15; 15190; English.

XX XX Sequences AAB23029-B23048 represent human SECX proteins. The SECX  
XX XX proteins and act as regulator of cellular proliferation and  
XX XX differentiation. SECX proteins or nucleotides are useful for diagnosing  
XX XX the presence of, or predisposition to, a disease associated with altered  
XX XX levels of SECX proteins and nucleotides. The SECX proteins are also  
XX XX useful to screen compounds that modulate SECX activity or expression. The  
XX XX interaction of a SECX protein with other cellular proteins may be useful  
XX XX to modulate the activity of a partner protein, cellular proliferation,  
XX XX cellular differentiation and cell survival. SECX nucleotides are useful  
XX XX for the recombinant expression of SECX protein, and may be used to detect  
XX XX SECX mRNA or genetic lesions in the SECX gene. They may also be used to  
XX XX modulate SECX expression (e.g., using antisense oligonucleotides). SECX  
XX XX nucleic acid sequences are also useful for identifying a cell or tissue  
XX XX type in a biological sample, and in forensic biology. SECX primers or  
XX XX probes are useful for detecting the presence of SECX nucleotides and for  
XX XX screening tissue cultures for contamination. Diseases that may be treated  
XX XX or prevented using SECX proteins or nucleotides include cancer (e.g.,  
XX XX colorectal carcinoma, prostate cancer), benign tumours, immune disorders  
XX XX (including autoimmune diseases, transient rejection, allergic disorders,  
XX XX infectious, inflammatory disorders, arthritis, haematopoietic disorders,  
XX XX skin disorders, cardiovascular disorders, atherosclerosis, restenosis,  
XX XX neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,  
XX XX surgical or traumatic wounds, spinal cord injury), and skeletal  
XX XX disorders.

XX XX Sequence 630 AA;

Qy 17 GAGPPESEPSISHGNYTKQYVFVGHKPGRNTTQPHRLDQYIMINMGLYIARCH 75  
|||  
Db 1 GSGFPPESEPSISHGNYTKQYVFVGHKPGRNTTQPHRLDQYIMINMGLYIARCH 60  
|||  
Qy 77 YTVYDITSHTEEVCSKKLTWKSRQADVTCTRMKGKHKDSCHNFIVLKKNDALFVCG 136  
|||  
Ub 61 YTVYDITSHTEEVCSKKLTWKSRQADVTCTRMKGKHKDSCHNFIVLKKNDALFVCG 120  
|||  
Qy 137 TNAPNSCRNYKMDTEPFQDEFSGMARCPYDAKHANVALFADGKLYSATVTOFLADAV 186  
|||  
Db 121 TNAPNSCRNYKMDTEPFQDEFSGMARCPYDAKHANVALFADGKLYSATVTOFLADAV 180  
|||  
Qy 197 IYASLGSBSPLTRVTKHDSKWKEPYFVQAVDYSDYIFFPREIAVENYTKGVVFFRVAQ 256  
|||  
Db 161 IYASLGSBSPLTRVTKHDSKWKEPYFVQAVDYSDYIFFPREIAVENYTKGVVFFRVAQ 240  
|||  
Qy 257 VCKNDXGSGSRVLEKQWTSFLKARLNCSVPDGSHEFYNLQAVTDVIRINGROWLATFS 316  
|||



AA094104;  
26-JUN-2001 (first entry).  
Human protein sequence SEQ ID NO:14328.  
Human: primer; detection; diagnosis; antisense therapy; gene therapy.  
Homo sapiens.  
EP1074617-A2.  
07-FEB-2001.  
28-JUL-2000; 2000EP-0116126.  
29-JUL-1999; 95JP-0248036.  
27-AUG-1999; 95JP-0300253.  
11-JAN-2000; 2000JP-0118776.  
02-MAY-2000; 2000CP-0183767.  
09-JUN-2000; 2000CP-0241895.  
(HELI-) HELIX RES INST.  
Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto C,  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
WPI; 2501-318749/34.  
Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs.  
Claim 8; SEQ ID 14328; 2537pp + CD ROM; English.  
The present invention describes primer sets for synthesizing 5602  
full-length cDNAs defined in the specification, where a primer set  
comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence and sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and  
in gene therapy. The primers are useful for synthesizing polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow chaining of the full-length  
cDNAs easily without any specialized methods. AAH03166 to AAH13628 and  
AAH13629 to AAH18742 represent human cDNA sequences; AAH02446 to  
AAH095893 represent human amino acid sequences; and AAH13629 to AAH13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.

589 SDSTAQEGYESRGMLDKHLLDSPDSTPGLGAVSSHHQDKGVRESYLKGHDLQVPV 648  
121 SDSTAQEGYESRGMLDKHLLDSPDSTPGLGAVSSHHQDKGVRESYLKGHDLQVPV 180  
649 TLLAIVILAFVAGVAFSGITVYCVCDHRKQKAVVQVQKEKETHSPRSMSSVTKLSGL 708  
181 TLLAIVILAFVAGVAFSGITVYCVCDHRKQKAVVQVQKEKETHSPRSMSSVTKLSGL 240  
709 FGDTQSKDKPKPEAILTPAXHNGKATPGNTAKYLKADQHHLLDALTPTPESTPTLOOKR 768  
241 FGDTQSKDKPKPEAILTPAXHNGKATPGNTAKYLKADQHHLLDALTPTPESTPTLOOKR 300  
769 KPSRGSEMERNONLINACTKMPYKSGSPVPTDLPURASPSHIPSVAVLPTIQOQYQHE 828  
301 KPSRGSEMERNONLINACTKMPYKSGSPVPTDLPURASPSHIPSVAVLPTIQOQYQHE 360  
829 YVQPKMSEVAQMALEDQAATLEYKTKHELSKSPNHNVENLDSLPKYPQREASL 888  
361 YVQPKMSEVAQMALEDQAATLEYKTKHELSKSPNHNVENLDSLPKYPQREASL 420  
889 GPTGASLSOTGLSKRLMEHSSSYGVYKESYPTNSLTRSHQATTIKRNYTNSNSSHLS 948  
421 GPTGASLSOTGLSKRLMEHSSSYGVYKESYPTNSLTRSHQATTIKRNYTNSNSSHLS 480  
949 RNOSFGRGNDPPAPQGVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKATPSLKP 1008  
481 RNOSFGRGNDPPAPQGVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKATPSLKP 540  
1009 DVPPKPSFAPLSTSMKFNDACT 1030  
541 DVPPKPSFAPLSTSMKFNDACT 562  
RESULT 10  
AAH03144  
ID AAH031444 standard; Protein; 562 AA.  
XX  
AC AAH031444;  
XX  
DT 06-NOV-2001 (first entry)  
XX Human polypeptide, SEQ ID NO: 3088.  
XX Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX Homo sapiens.  
XX  
XX EP:130094-A2.  
XX 05-SEP-2001.  
XX 07-JUL-2000; 2000EP-0114389.  
XX 08-JUL-1999; 99CP-0194486.  
XX 11-JAN-2000; 2000CP-019774.  
XX 02-MAY-2000; 2000CP-0183765.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;  
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI; 2001-524255/58.  
XX N PSDS; AAH04365.  
XX  
XX #30 Primers useful for synthesizing full length cDNA clones and their  
XX use in genetic manipulation.  
XX  
XX Claim 8; SEQ ID NO 3088; 1380pp + sequence listing; English.  
XX  
XX The invention relates to primers for synthesizing full length cDNA  
XX clones, 830 cDNA molecules encoding a human protein have been  
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

Query Match 54.2%; Score 2952; DB 22; Length 562;  
Best Local Similarity 100.0%; Freq. No. 1,40-247;  
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0.  
469 MSVYASEKSYGVVEDKRMVQCLDRASSLLYAFSTCVKYPAGCRPHGRKCKTCAS 525  
1 MSVYASEKSYGVVEDKRMVQCLDRASSLLYAFSTCVKYPAGCRPHGRKCKTCAS 50  
529 RDPYGVWKEGACSHLSNSRLTTEQDTERGNTDGLDCNHSFVALNGHSSSLPSTT 586  
61 RDPYGVWKEGACSHLSNSRLTTEQDTERGNTDGLDCNHSFVALNGHSSSLPSTT 120



670 VYVCDHRRKDAVYVORKEKELTISREGSYSVTK:SGLEGTQSKDKPKFAILTPLMHN 729  
|||||  
503 VYVCDHRRKDAVYVORKEKELTISREGSYSVTK:SGLEGTQSKDKPKFAILTPLMHN 562  
|||||  
730 GKLATPGNTAKMLIKACCHHLDJALPTPESTPTCCQKEKTFEGSGREMERNCNLINACTK 739  
|||||  
563 GKLATPGNTAKMLIKACCHHLDJALPTPESTPTCCQKEKTFEGSGREMERNCNLINACTK 522  
|||||  
790 DNPFGSPVITDPLRASPDSHIFSVAVPIIQOQYQYQYVQPKYSEVACMALEDQAT 849  
|||||  
623 DNPFGSPVITDPLRASPDSHIFSVAVPIIQOQYQYQYVQPKYSEVACMALEDQAT 582  
|||||  
950 LEYKTIKEHLSSKSPNHNJLVENLQSLPKYVQPEASLGPFGASLSQTGSKRLMEHHS 909  
|||||  
683 LEYKTIKEHLSSKSPNHNJLVENLQSLPKYVQPEASLGPFGASLSQTGSKRLMEHHS 742  
|||||  
910 SSYGYEYKSYPTNSLTRSHQATILKRNNTNSNSH:SPNQSGRGDNPPAPQVLSI 969  
|||||  
743 SSYGYEYKSYPTNSLTRSHQATILKRNNTNSNSH:SPNQSGRGDNPPAPQVLSI 802  
|||||  
970 QVHSSQPSQAVTVSGQPELKNYNSLTRSGLXRTSLYDFDPEKSPAPLSTSKYKPNAD 1029  
|||||  
923 QVHSSQPSQAVTVSGQPELKNYNSLTRSGLXRTSLYDFDPEKSPAPLSTSKYKPNAD 862  
|||||  
1030 T 1030  
863 T 863  
RESULT 12  
ABG04620 standard; Protein; 962 AA.  
ABG04620;  
13-FEB-2002 (first entry;  
Novel human diagnostic protein #4611.  
Human; chromosome mapping; gene mapping; gene therapy; forensic;  
food supplement; medical imaging; diagnostic; genetic disorder.  
Homo sapiens.  
WO250175067-A2.  
11-OCT-2001.  
3C-VAR-2001; 2001WO-US03651.  
31-MAR-2000; 2000US-0542217.  
23-AUG-2000; 2000US-0649167.  
(HYSE-) HYSEQ INC.  
Drmanac RT, Liu C, Tang YT;  
WPI; 2001-639162/73.  
N-P5DB; AAS68807.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -  
Claim 20; SEQ ID No 14979; 103pp; English.  
The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and for chromosome  
mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags

for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. ABG04620-ASG10377 represent novel human  
diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from Wipo  
at ftp.wipo.int/pub/published\_pat\_sequences.

Sequence 962 AA;  
Query Match 52.83; Score 2879.5; DB 22; Length 962;  
Post-Local Similarity 70.33; Pred. No. 7.1e-241;  
Matches 593; Conservative 30; Mismatches 122; Indels 101; Gaps 12;  
QY 147 YKQDT-DEPFGEFSGMARQPYDAKHANVALFA-----DGKLYSATVTCDFLAD 194  
|||  
DE 95 YEDTWLCKWSSSWTLVFTHPFSAALHSAITARSUTGSLTSPDGKLYSATVTCDFLAD 154  
|||  
QY 195 AVYRSLGESPTLRTYKHSKWLKEPYFVQAVQYGDYYPFFREIAVEYNTYKVVPRV 254  
|||  
DE 155 AVYRSLGESPTLRTYKHSKWLKEPYFVQAVQYGDYYPFFREIAVEYNTYKVVPRV 214  
|||  
QY 255 AQVCKNDMGSGQVLEKQWTSF:KARLNCVSPGDSHFYFNILQAVTDVIRGQVYLAT 314  
|||  
DE 215 AQVCKNDMGSGQVLEKQWTSF:KARLNCVSPGDSHFYFNILQAVTDVIRGQVYLAT 274  
|||  
QY 315 PSTPYNSIPGSVAVCAVEXLDIASVFTGRFKEKSPDSTWTPVDERVPPKPRGCCAGSSS 374  
|||  
DE 275 PSTPYNSIPGSVAVCAVEXLDIASVFTGRFKEKSPDSTWTPVDERVPPKPRGCCAGSSS 334  
|||  
QY 375 LERYATSNPEPDJLNFIKTHPLMDSAVESIFRNPFLRTNVRJRJTKIAVDTAAAGPYQN 434  
|||  
DE 335 LERYATSNPEPDJLNFIKTHPLMDSAVESIFRNPFLRTNVRJRJTKIAVDTAAAGPYQN 394  
|||  
QY 435 HTVYFLGSEKGIILKFLARIGNSGFNDSLFLBEMSYNSEKSYDGVEDKRMGVQDLR 494  
|||  
DE 395 HTVYFLGSEKGIILKFLARIGNSGFNDSLFLBEMSYNSEKSYDGVEDKRMGVQDLR 454  
|||  
QY 495 ASSSLYVAFSTCVIKVPLGRCERHGHCKKTCIASRDPYCGWIKEGGACSHLSPNSRLT-- 552  
|||  
DE 455 ASSSLYVAFSTCVIKVPLGRCERHGHCKKTCIASRDPYCGWIKEGGACSHLSPNSRLTORG 514  
|||  
QY 553 FEQDIERGNTDGLGDC-----HNSFVAJNGHSSS----- 591  
|||  
DE 515 FVWFVVTGST--LHPCILLTSLWQTHPSY-----HSSSETGRPVKXGGERGKCTLNHFEL 567  
|||  
QY 582 --LPSTTTTSCSTAQEGYESRGCM-----LWKGHLSDSPDTPLGAVSSHNHCKKGVII 634  
|||  
DE 563 GYKLFVSYTLLRVVIGVGLKRCOTLFFRRRLRW-----PQGRRRKRSQLEAQRV-- 616  
|||  
QY 635 RESYLKGHCDQVPTLLA-AVILAFVMGAVFSGITVYVCDHERKDVAVVQREKELTHS 694  
|||  
DE 617 RESYLKGHCDQVPTLLA-AVILAFVMGAVFSGITVYVCDHERKDVAVVQREKELTHS 676  
|||  
QY 695 RRGSMSVTKLSGLFDQTSKQPKPEALITPLMHNKGLATPGNTAKMLIKACCHHLDJLA 754  
|||  
DE 677 RRGSMSVTKLSGLFDQTSKQPKPEALITPLMHNKGLATPGNTAKMLIKACCHHLDJLA 736  
|||  
QY 755 LPTPESTPLQOKRKPSRSGREMERNCNLINACTKXPMGSPVITDPLRASPDSHIPS 814  
|||  
DE 737 LPTPESTPLQOKRKPSRSGREMERNCNLINACTKXPMGSPVITDPLRASPDSHIPS 796  
|||  
QY 915 VVVLPIITCCQYQYQYVQPKYSEVACMALEDQATLEYKTIKEHLSSKSPNHNJLVENL 974  
|||

D 797 HNSCPSSRATSMSTWSPXAXWFWWRKTRPFWHSIFPSNINSAARVPTNEXILKRW 856  
Y 875 DSLUPPK-----VFQREASJGPPGAS-----LSQTLSELEMHSSSYGVYKRSY 920  
D 657 TACPPKPHSRPPWPREPCJRFVXASGKCTPPLTLTIZ----- 999  
Y 921 PINSLTRSHQAT 932  
b 900 GATPRTSRHEAT 911  
  
RESULT 13  
AB923688  
D AAB92688 standard; Protein; 507 AA.  
X  
X AAB92688:  
X 26-JUN-2001 (first entry)  
X Human protein sequence SEQ ID NO:11073.  
E Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
W Homo sapiens.  
S  
S EPI074617-A2.  
X  
X 07-FEB-2001.  
X 26-JUL-2000; 2000EP-016126.  
X 29-JUL-1999; 90JP-0348236.  
X 27-AUG-1999; 90JP-0302253.  
R 11-JAN-2000; 2000JP-0118776.  
R 02-MAY-2000; 2000JP-0183767.  
R 09-JUN-2000; 2000JP-0241939.  
X (HELI-1) HELIX RES INST.  
X Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto G,  
I Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T,  
WPI; 2001-318745/34.  
X  
X Primer sets for synthesizing polynucleotides, particularly the 5602  
T full-length cDNAs defined in the specification, and for the detection  
T and/or diagnosis of the abnormality of the proteins encoded by the  
T full-length cDNAs -  
T  
T Claim 8; SEQ ID 11073; 2537pp + 1D PCM; English.  
S  
S The present invention describes primer sets for synthesizing 5602  
C full-length cDNAs defined in the specification, where a primer set  
C comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
C to the complementary strand of a polynucleotide which comprises one of  
C the 5602 nucleotide sequences defined in the specification, where the  
C oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
C of an oligonucleotide comprising a sequence complementary to the  
C complementary strand of a polynucleotide which comprises a 5'-end  
C sequence and an oligonucleotide comprising a sequence complementary to a  
C polynucleotide which comprises a 3'-end sequence, where the  
C oligonucleotide comprises at least 15 nucleotides and the combination  
C of the 5'-end sequence/3'-end sequence is selected from those defined in  
C the specification. The primer sets can be used in antisense therapy and  
C in gene therapy. The primers are useful for synthesizing polynucleotides,  
C particularly full-length cDNAs. The primers are also useful for the  
C detection and/or diagnosis of the abnormality of the proteins encoded by  
C the full-length cDNAs. The primers allow obtaining of the full-length  
C cDNAs easily without any special used methods. AAH0166 to AAH1628 and  
C AAH1363 to AAH1542 represent human cDNA sequences; AAB92446 to  
C AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
C represent oligonucleotides, all of which are used in the exemplification  
C of the present invention.

XX Sequence 507 AA:  
SQ  
Query Match 45.1%; Score 2621.5; DB 22; Length 507;  
Best Local Similarity 95.2%; Pred. No. 7e-219;  
Matches 507; Conservative 0; Mismatches 0; Indels 55; Gaps 1;  
  
QY 469 MSVYNSEKSYDGVEDKRINQVLDRASSSLYVAFSTCVIKVPLGRCEHKGKCKTCIAS 528  
DB 1 MSVYNSEKSYDGVEDKRINQVLDRASSSLYVAFSTCVIKVPLGRCEHKGKCKTCIAS 60  
  
QY 529 RDPYCGWIKEGGASHLSPNSRLTFEQDIBRGTDGLGCHNSFVALNGHSSSLPSTTT 588  
DB 61 RDPYCGWIKEGGASHLSPNSRLTFEQDIBRGTDGLGCHNSFVALN----- 108  
  
QY 589 SDSTAQEGYESRGMDLWKELLDSPDSTPLGAVSGHHNQCKKGVIRESYKLGHDLPV 648  
DB 109 -----GVRESYKLGHDLPV 125  
  
QY 649 TLLAIAVILAFVMAVESGITVYVCVCHRRKDVAVVQVQKEKELTHSRGSMSSVTKLSGL 708  
DB 126 TLLAIAVILAFVMAVESGITVYVCVCHRRKDVAVVQVQKEKELTHSRGSMSSVTKLSGL 185  
  
QY 709 PGDTOSKDPKPEALITPLMENGKLPATPGNTAKVJIKADQHHLDLTAIPTPSTPTLOQKR 768  
DB 186 PGDTOSKDPKPEALITPLMENGKLPATPGNTAKVJIKADQHHLDLTAIPTPSTPTLOQKR 245  
  
QY 769 KPSRGSREWRNQNLIKACTKDVPPMGSPVITDPLRASPSHPSVVLPIPTQQGYCHE 828  
DB 246 KPSRGSREWRNQNLIKACTKDVPPMGSPVITDPLRASPSHPSVVLPIPTQQGYCHE 925  
  
QY 829 YVDPQKYSVAQVAKALEDCAAALEYKTKKEHLSSKSPNAGNLYENLDLSPKVPQREASL 888  
DB 306 YVDPQKYSVAQVAKALEDCAAALEYKTKKEHLSSKSPNAGNLYENLDLSPKVPQREASL 948  
  
QY 889 GPPGASLSQTLSELEMHSSSYGVYKSYPTNSLTPSHQATTAKRNTNNSNSHLS 948  
DB 366 GPPGASLSQTLSELEMHSSSYGVYKSYPTNSLTPSHQATTAKRNTNNSNSHLS 925  
  
QY 949 ENQFGRGDNFPFAPQVDSIQVHSSQPSQCAVTVSOPSLNAYNSLTRSGLKRTPSLKP 1006  
DB 426 ENQFGRGDNFPFAPQVDSIQVHSSQPSQCAVTVSOPSLNAYNSLTRSGLKRTPSLKP 485  
  
QY 1009 DVPPKPSFAPJSTSMKENDACT 1030  
DB 486 DVPPKPSFAPJSTSMKENDACT 507  
  
RESIDU 14  
AAG63213  
ID AAG63213 standard; Protein; 1086 AA.  
XX  
AC AAG63213:  
XX  
DT 01-OCT-2001 (first entry)  
D2 Amino acid sequence of a human semaphorin-like polypeptide.  
XX  
KW Expressed sequence tag; R57; semaphorin-like protein; neuronal growth;  
KW spinal cord damage; neurodegenerative disease; genetic neuronal defect;  
KW immunological disorder; lymphocyte dysfunction; viral infection; cancer.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..16  
FT /note= "signal peptide"  
FT Domain 671..693  
FT /note= "transmembrane domain"  
XX WC20C153466-A1.  
XX 26-JUL-2001.  
PD



22-DEC-2000; 2003KO-US33158.  
 23-DEC-1999; 99US-0471275.  
 21-JAN-2000; 2000US-0483725.  
 25-APR-2000; 2000US-0552317.  
 31-AUG-2000; 2000US-0653274.  
 (HYSF-) HYSEQ INC.  
 Boyle BC, Yeung CY, Atterburn VC, Nize NK, Tang YC, Liu C.  
 Dimaac RT;  
 WPI; 2001-476114/51.  
 N-PSDB; AA42197.  
 New polynucleotides and the encoded polypeptides having human secreted semaphorin-like polypeptide, useful in treating neurodegenerative diseases, viral infections or cancers, or diagnosing and mapping genetic neuronal defects  
 Claim 12: Page 124-127; 142pp; English.  
 The present sequence represents a human semaphorin-like polynucleotide. Semaphorin-like polynucleotides and polypeptides are useful in therapeutic, diagnostic and research methods. These are particularly useful in modulating neuronal growth regenerative capacity (e.g. in the case of spinal cord damage), treating neurodegenerative diseases, diagnosing and mapping genetic neuronal defects. These are also useful in treating immunological disorders arising from T and B lymphocyte dysfunction, or in treating viral infections or cancers.  
 Sequence 1086 AA:  
 Query Match 43.1%; Score 2381.5; DB 22; Length 1086;  
 Best Local Similarity 45.5%; Pred. No. 1.1e-195;  
 Matches 512; Conservative 159; X.matches 306; Indels 147; Gaps 30;  
 1 MRSEALLVFTLL---HFAGAGFFEDSDSPISISHGNYTKQYVFGVGHKPGNTTORHLD 57  
 1 MRVFLLCAYLLVMVSQLRAVFFEDSDSPNTVYHSGQVVFVFG-RPNSGSO-HRLD 58  
 58 IQVIMMGTEVIAARDHYTVDTTSHTTEEYQSKLTKSRQADVDTCRKKGKHDEC 117  
 59 PQMKIKRQTSIAGRDQYVTKLMEPKTEVTPKKLTWRSRQDRENCVAKGKHDEC 118  
 118 HRFIKVLLKNDALFVOTNAPNSCRNKKYKDTLEPFGQEPSHMACPYQAKHANALP 177  
 119 HRFIKVPRNDEKVFVCGTNAPNPKYRYSLTLEYEGEEISGLARCFPDARQTNALP 178  
 178 ADGKYSATVTPFLALDAVYRSLGSPTLRVRKIDSKWLZEPFFVQAVDYCHYFFTER 237  
 179 ADGKYSATVADFLASDAVYRSMGDGSLRTKYDSMKIKEPHEFLHAEYGNYYFFTER 238  
 238 ELAVENYTKGVYFVRAVQCKNMGSGRVLEKQWTSFLKAPLNCVPGSGHVFENLIQ 297  
 239 ELAVENNLKAVYRVARICKNMGSGRVLEKWTSTFLKARLNCVPGGFFPFDLIQ 298  
 298 AVTDVIRNGRWALATSTPTYSNIFGSAVQVMDLDAVFTGFEKQKSPSTWTFVP 350  
 299 SITDIQNGITVGVFTQLNSIFGSAVCAFNMDELEKVGKGFKEQKTPDSWTVAP 359  
 358 DERVPKPRGCGAGSSLEHYATSNEPCTINFTKTHPLKZEVNFSINRWFEETVVP 417  
 359 EDKVPKPRGCGAKGLADAVYATSIDPDETASF-KSHPLMDGSAVPIADEPWFKTAVR 418  
 418 YRLTKIADTAGPYQNHVTFVLSGKGLKFLASLNSGFLNDSLEFENYSWYSEK 476  
 419 YRLTALSDHSGPYQNYTVTFVSGAGVGLVLAKE--TSFPLNDSVLEDEIENYNAK 477  
 477 CSYDGVEDHRIIMNQICRASSSLYVAFSTCVIKVPLGRCKRGCKKTKCIASPRPYCGWI 516  
 477 CSAENEEDKKVISLQDKDHHALYVAFSSCIIPFLSRCEHYGSKKSCIASRDPYCCWL 516

QY 537 KEGCAASHLSPNSLIT-----FEDIERGNTDGLGDOHNSFVALNGHSSSL 582  
 DB 537 SQ-GSCORVTPGMJLTEDFFAFHNHSAEGYQDTEFGNTAHLGCOH-----EI 584  
 QY 583 LPSTTTSDSTAQCEVESRGO-----MLD-WKHLA-----610  
 DB 585 LPTSTTPD-----YKIFGQPTSDMEVSSSVTTMASIPEITPKVIDTWPKLSRKKFV 638  
 QY 611 --DSP---DSTPLGAVSSHHQCKKGVRESYKGDHQLVPVTLAIAVILAFYGVAVF 665  
 DB 639 VQDQENTSDFTPLSGI-----PKGVMEVQSGESNQVHMNVLTTCVFAAFVLGAPI 691  
 QY 666 SGITVTCVCD-HERKQVAVYORKEKELTHSRSSMSVTKLSQLFG-----DTQSKDKPKPE 720  
 DB 692 AGVAVCYRDMFYKARKI--HKDAESAQCTDSSGSFAKLNGLFDSPVKEYQDN:DSPX 749  
 QY 721 ALLTELHNGKLAFTGNITAKVLKADQHHLDDLALPTPESTPLQQRKFSRGSREKERN 780  
 DB 750 LYNNLLTSRKELPPNCTKSNVMDHRCQPPDLALPTPESTPLVLRQKTLQAMKSHSEKAH 859  
 QY 781 QNLINACTKMPMGSEVITPLPLRASFSHPVVVLP:TCQGYOHEY-----929  
 DB 810 GH-GASRKEITQFFSSPP2HSL--SHGHIPSAIVLPNATHDYNTSFNSNAKAEKK 865  
 QY 830 ---VDCP---KMSEVACMALEDQAA:LEYKTIYEHLSKSPN-----HGYNLVE 872  
 DB 866 LQNDHFLTYSKSKREHRAVSERNIT--NLLKHLNDPNKSNKAINGE:QMAHOKMLD 923  
 QY 873 NLDSJ---PKVQCRFASLQPPGASLSQTLKSRLEMHSSSY---GVNYSKRYPTNLSI 926  
 DB 924 PKGSMSEVPKVENREASLYSFFSTLPRKSPTRKVDVPTTPGVMTSLERQRYHKNSSC 983  
 QY 927 RSHQATTLKANTNSNSHLSRQSGFRGQ-NPPAPQAVDS:QVRSQFSCQAVTVSR 985  
 DB 984 R-HSISAMPK-NLNSPGLVLSRQPSMNRGYPMTTGAKVVDVIQ-----GTPSVHL 1034  
 QY 986 QPSLNAYNSLT-----RSQKDTPLSKPDPVPKPSFAPLSTSKMP 1025  
 DB 1035 QPSLGRSSYTSNGT:FRTGKRTPLSKPDPVPKPSFVQTPSVRP 1086  
 RESULT 15  
 AEG79177  
 ID AEG79177 standard; Protein; 1088 AA.  
 AC AEG79177:  
 XX 15-NOV-2002 (first entry)  
 DE Human semaphorin-like protein #5.  
 DE Human; NCNV; neurodegenerative disease; Alzheimer's disease; anxiety;  
 KW Parkinson's disease; Huntington's disease; neurological disorder;  
 KW schizophrenia; manic depression; mental retardation; angina pectoris;  
 KW cardiovascular disease; acute heart failure; myocardial infarction;  
 KW muscular disease; muscular disorder; retinal disease; photoreception;  
 KW gastritis; keratinisation disorder; cancer; ovarian cancer; melanoma;  
 KW immunological disorder; inflammatory disease; immune disease; diabetes;  
 KW bacterial infection; fungal infection; protozoal infection; obesity;  
 KW viral infection; reproductive system disorder; metabolic disturbance;  
 KW anorexia; wasting disorder; chronic disease; infectious disease;  
 KW dyslipidaemia; cub; sushi; myelin; von willebrand factor; kielin;  
 KW semaphorin; serine/threonine protein kinase; TGF-beta binding;  
 KW mas proto-oncogene; ribonuclease pancreatic precursor; aminotransferase;  
 KW colloid-like 2; cysteine sulfinic acid decarboxylase.  
 OS Homo sapiens.  
 PK WO200264791-A2.  
 PD 22-AUG-2002.  
 XX

F 10-DEC-2001; 2001WO-US#8369.

X 08-DEC-2000; 2000US-254327P.

R 14-DEC-2000; 2000US-255648P.

R 15-MAY-2001; 2001US-29-037P.

R 08-JUN-2001; 2001US-297153P.

R 08-JUN-2001; 2001US-309258P.

R 29-AUG-2001; 2001US-315639P.

R 01-OCT-2001; 2001US-326393P.

X (CURA-) CURAGEN CORP.

X I Alsobrook JP, Anderson DW, Burgess CE, Bolodog FL, Casman SJ,

I Colman SD, Edinger SR, Eilerman K, Gerlach V, Gorman J, Grasse WM,

I Guo X, Herrmann J, Keuda R, Lopley DM, Li L, MacDougall JR,

I Millet I, Pena CE, Peyman CA, Rastelli L, Rieger DK, Shinkate RA,

I Smithson G, Spytex KA, Stone MJ, Tchertev VT, Vernet CA, Voss E2,

I Zorhusen SD, Zhong H, Zhong M;

X WPI; 2002-643486/69.

R N-PSDB: ABS64384.

X New NOVX polypeptides and polynucleotides useful for treating or

X preventing e.g. neurodegenerative diseases, neurological disorders,

X cardiovascular diseases, muscular diseases and disorders, or

X immunological diseases

X Claim 1; Page 56; 299pp; English.

X The present invention relates to new NOVX polypeptides. The polypeptides,

X polynucleotides and antibodies are useful in the manufacture of a

X medicament for treating or preventing neurodegenerative diseases (e.g.

X Alzheimer's disease, Parkinson's disease, or Huntington's disease),

X neurological disorders (e.g. anxiety, schizophrenia, manic depression or

X mental retardation), cardiovascular disease (e.g. acute heart failure,

X angina pectoris or myocardial infarction), muscular diseases and

X disorders, retinal diseases (including those involving photoreception,

X deafness and keratinisation disorders), cancer (e.g. ovarian cancer or

X melanoma), immunological disorders, inflammatory and immune diseases,

X bacterial, fungal, protozoal and viral infections, and reproductive

X system disorders. The proteins of the invention may be used to screen

X drugs or compounds that modulate the NOVX protein activity or expression,

X as well as to treat disorders characterised by insufficient or excessive

X production of NOVX protein or protein forms that have decreased or

X aberrant activity compared to NOVX wild type protein, such as diabetes,

X obesity, metabolic disturbances associated with obesity, such as diabetes,

X wasting disorders associated with chronic diseases and various cancers,

X infectious diseases and various dyslipidaemias. The nucleic acid

X sequences of the invention may be used in chromosome mapping.

X Identifying an individual from mouse biological samples (tissue typing).

X and in forensic identification of a biological sample. The present

X amino acid sequence represents a NOVX protein of the invention.

X X Sequence 1288 AA;

X Query Match 43.14; Score 2349.5; DB 23; Length 1088;

X Best Local Similarity 45.43; Pred. No. 122e-154;

X Matches 512; Conservative 156; Mismatches 309; Indels 149; Gaps 31;

Y 1 MRSEALLLYFELL---HFAGAGPFDESEPSI---SHGNY-KQYVVFVGHKPGNTTQPHR 45

b 1 MRVFLCAYILLMVLSQRAVSPFEDDEPLNTVDYHCKSSRQYVFRG-RFSNESQ-HR 58

Y 56 LDQVIMIMKGLVIARDHYITVDISTHEEYCSKLTWKSRAQVDVTCNKQKHKD 115

b 59 LDFQLMLKIRETLVIAGRDQVITVNLNEXPEKTEY-WQCKTWESRQCDRENCAYKQKHKD 118

Y 116 ECHNFIVJLKKNDCA-FVGTGNFNSCRVYKVDLTLEPGDSSGMARCPYKAKANVA 125

b 119 ECHNFIVFVRNDEMVFVCTGNFNPNCYRYRYVSTLEYGCEI-SGLIARCPFDARQTNVA 128

Y 176 LFADGKLYSATVDFLAIDAVIYNLSGESPTLRTVKHDSKMLKEPFPVQAVYGVYVFF 235

DB 179 LFADGKLYSATVDFLAIDAVIYNLSGESPTLRTVKHDSKMLKEPFPVQAVYGVYVFF 238

QY 236 PREAVENYINMGKYVFPFVACVCKNDMGSCORVLEKQWTSFLLKARLNCSPVGSHEFYFNI 295

DB 239 PREAVENHNLGRAVSRVARICKNDMGSCORVLEKHTWTSFLLKARLNCSPVGSFYPFDV 298

QY 296 LOAVTDVIRINGRDVJATFSTPYNSIPSSAVCAKMLDIASVETGRPRECKSPDSTWTP 355

DB 299 LOSITDIQINGFTVVGVTFTQNSIPGSAVCAFMDDIEKVFGRKECKTPDPSVNTA 358

QY 356 VPDERVPKPRPGCCAGSSLSERVATSNFEDDTLNFPIKTHPLMDEAVPSTFNRPWFRTM 415

DB 359 VPEKVPKPRPGCCCKKGLAEAYKTSIDFDETLSP-KSHPLMDSAVPPADPEWFKTR 418

QY 416 VRYLTIAVDYTAAGPYQNHVTVFLGSEKGIILFLARIGNSGF-LNDSJFLBMSVYNS 474

DB 419 VYRLTAISVDHSAGPYQNTVTFVSGAGWLKVLAK--TSPFSLNDSVLEIEIAYNH 476

QY 475 EKCSYDGVEXKRIKMGMDLRASSLYVAFSTCVIKVPLGHCERHGKCKTCTASRDYCG 534

DB 477 AKCSABNEEDKXVLSLO-DKHDAIYVAFSSCIIRPLSRCEYRGSCCKSCIASRDYCG 536

QY 535 WIKEGGACSHLSPNSRLT-----PEQDIERGNTDGLGCHNSFVALNCHSS 580

DB 537 WLSQ-GSGGAVTPGMULLTEDPFAFHNSABGYEQDTEFCNTAHLGCKH----- 584

QY 581 SJLPSTTTSSTAGEYSESQG-----YKIFGFTSDMEVSSSVYTWASIFE-TPKVIDTWRPKLTSSRK 610

DB 585 EILPSTSTPD-----DSTPLGAVSSHCHCKGVIRSY-KGHQQLVFTLLAIVILAFVYGA 638

QY 611-----DSP-----PKGVWEVQSGESNQVMHVVLTCTVPAAFVLGA 663

DB 639 FVWQDDNTSFTDPLSGI-----HKAESAQSCTDSSGSFAKINGLPSVKEYVQQNIDS 691

QY 664 VFSGITVYVCCD-HRRKDVAVQPKKELTHSRPGSMSSVTKLSGFS-----DTQSKDKPK 718

DB 662 FIAGVAVCYRDMVVRKRI--HKDAESAQSCTDSSGSFAKINGLPSVKEYVQQNIDS 749

QY 719 PERALTPLMNGKLTATNTAKMLKADQHHLDLTALPTPESFTLQOKKPKSGSREWE 778

DB 750 PKLYSNLTLSKELPPKGDTSKVMYMRHGPPPELAALPTPESTFVLHQKTLQAMKSHSEK 809

QY 779 RNOQLINACTKQVPMCSFPIPTDLPURASPSHPSVWVLPITCCQGYQHEV----- 829

DB 810 AFQH--GASRKEETCFPFSSPPPHSPL--SHGHIPSAIVLPLKATHDNTSFSNNAKAE 865

QY 830-----VQCP---KMEFVAQWALDQAAATLEYKTIKEHLSSKSPN-----HGVNL 870

DB 866 KKLQMLCHPLTKSSKXKHRRSVDNRNTD--NQLLKLNDPNSNPKAIMEDIQWAFQNLN 923

QY 871 VENLDST---PFKYPQREASLGPFGASLCTGLSKRLEMHSSSY---GVYKESYFNS 924

DB 924 LCPMGSSEVEPKYFNREASLYSPSTLPRKSPFKRVFVPTTPGVMTSJRGRGYHNS 963

QY 925 LTRSHQATTLKRNNTKSNSSHLSRNSQSGRGC-NPQPAQRVDSIQVHSSQSGCAVTV 963

DB 984 SCR-HS-SAMKP-NLNSPQVGLSRQPSMRNGCHMPTPTGAKVDYIQ-----GTPSV 1034

QY 984 SRQPSLNAYNSLT-----ESGLKETPSLKPDPVPPKPSFAPLSTSMKP 1025

DB 1035 HLQPSLSRQSSYTGTLPRGLKSTPSLKPDPVPPKPSFAPLSTSMKP 1082

Search completed: October 23, 2003, 17:10:34  
Job time : 56 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
M protein - protein search, using sw model  
Un on: October 23, 2003, 17:10:42 : Search time 64 Seconds  
(without alignment)  
2533.156 Million ch1 updates/sec  
Title: US-09-856-681-2  
Perfect score: 5450  
Sequence: 1 MRSEALLLVFLLHFAGAGF.....PPKPSFAPLSMKPNDACT 1030  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 629382 seqs, 167460630 residues  
Total number of hits satisfying chosen parameters: 629382  
Minimum DB seq length: 3  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC7\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PC7US\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US03\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US0C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US0 NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY-ES				Query		Description	
Result No.	Score	Match	Length	DB ID			
1	5422.5	99.5	1047	11	US-09-957-187-85	Sequence 85, Appl	
2	492	90.3	939	11	US-09-957-187-4	Sequence 4, Appl	
3	4590.5	94.2	884	11	US-09-957-187-6	Sequence 6, Appl	
4	3373	61.9	630	11	US-09-957-187-30	Sequence 30, Appl	
5	3362	61.7	626	11	US-09-957-187-3	Sequence 3, Appl	
6	2013.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
7	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
8	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
9	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
10	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
11	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
12	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
13	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
14	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	
15	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl	

16	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl
17	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, Appl
18	2009.5	36.9	888	12	US-10-141-698-544	Sequence 544, Appl
19	2009.5	36.9	888	12	US-10-141-702-544	Sequence 544, Appl
20	2009.5	36.9	888	12	US-10-141-704-544	Sequence 544, Appl
21	2009.5	36.9	888	12	US-10-142-421-544	Sequence 544, Appl
22	2009.5	36.9	888	12	US-10-142-432-544	Sequence 544, Appl
23	2009.5	36.9	888	12	US-10-142-767-544	Sequence 544, Appl
24	2009.5	36.9	888	12	US-10-142-767-544	Sequence 544, Appl
25	2009.5	36.9	888	12	US-10-143-033-544	Sequence 544, Appl
26	2009.5	36.9	888	12	US-10-144-994-544	Sequence 544, Appl
27	2009.5	36.9	888	12	US-10-145-631-544	Sequence 544, Appl
28	2009.5	36.9	888	12	US-10-145-633-544	Sequence 544, Appl
29	2009.5	36.9	888	12	US-10-145-748-544	Sequence 544, Appl
30	2009.5	36.9	888	12	US-10-145-748-544	Sequence 544, Appl
31	2009.5	36.9	888	12	US-10-145-823-544	Sequence 544, Appl
32	2009.5	36.9	888	12	US-10-145-826-544	Sequence 544, Appl
33	2009.5	36.9	888	12	US-10-145-870-544	Sequence 544, Appl
34	2009.5	36.9	888	12	US-10-145-876-544	Sequence 544, Appl
35	2009.5	36.9	888	12	US-10-145-959-544	Sequence 544, Appl
36	2009.5	36.9	888	12	US-10-146-724-544	Sequence 544, Appl
37	2009.5	36.9	888	12	US-10-146-725-544	Sequence 544, Appl
38	2009.5	36.9	888	12	US-10-146-795-544	Sequence 544, Appl
39	2009.5	36.9	888	12	US-10-147-495-544	Sequence 544, Appl
40	2009.5	36.9	888	12	US-10-147-501-544	Sequence 544, Appl
41	2009.5	36.9	888	12	US-10-147-504-544	Sequence 544, Appl
42	2009.5	36.9	888	12	US-10-147-506-544	Sequence 544, Appl
43	2009.5	36.9	888	12	US-10-147-509-544	Sequence 544, Appl
44	2009.5	36.9	888	12	US-10-147-510-544	Sequence 544, Appl
45	2009.5	36.9	888	12	US-10-147-511-544	Sequence 544, Appl

ALIGNMENTS

RESULT 1  
US-09-957-187-85  
Sequence 85, Application US/09957187  
Publication No. US20030054514A1  
GENERAL INFORMATION:  
APPLICANT: Sharklets, Richard A.  
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND PROTEINS ENCODED THEREBY  
FILE REFERENCE: 15966-540 CIP  
CURRENT APPLICATION NUMBER: US/09/957,187  
CURRENT FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/123,667  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/520,721  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 65/234,582  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 60/233,798  
PRIOR FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/174,485  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: Patentin ver. 2.1  
SEQ ID NO 85  
LENGTH: 1047  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-957-187-85  
Query Match: 99.5%, Score 5422.5, DB 11, Length 1047;  
Best Local Similarity: 98.3%, Pred. No. 0;  
Matches 1029, Conservative: 0, Mismatches: 1; Indels: 17, Gaps: 1;  
CY 1 MRSEALLLVFLLHFAGAGFPEDESEISISGNYTKQYFVGHKPGKGRNTTORHLDIQM 60  
DB 1 MRSEALLLVFLLHFAGAGFPEDESEISISGNYTKQYFVGHKPGKGRNTTORHLDIQM 60  
CY 61 IMINXGTVYIARCHIIVTD:DTSH:EEIYCSKKLTKWSRCAOVTCRMKXKHKDECHNF 120

```
61 1NVMNGTVAARHITVVIDISHTETVCSKLTWKSACADCTCRMGKKGKSCHE 120
121 1KVLKXGCEALFYCGTNAFNSFENYKMTLEFFDFSSVACQPYDAKHANVALFAQ 140
121 1KVLKXGCEALFYCGTNAFNSFENYKMTLEFFDFSSVACQPYDAKHANVALFAQ 140
181 1KYSATVTDFLAIDAVIYRS:GESPTLRTVKHDSKWKKEPYFQAVDGYIYFFREIA 240
181 1KYSATVTDFLAIDAVIYRS:GESPTLRTVKHDSKWKKEPYFQAVDGYIYFFREIA 240
241 1VEYNTMGKVVFPVAVQVCKXDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFYENILQAVT 300
241 1VEYNTMGKVVFPVAVQVCKXDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFYENILQAVT 300
301 1DVIRNGRDVVLATFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEKQSPSTWTFVPCR 360
301 1DVIRNGRDVVLATFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEKQSPSTWTFVPCR 360
361 1VPRPCCGAGSSSLERYATSNFDDTLNFIKTHPLMDSVPS:FNRPFLSTWYAYRL 420
361 1VPRPCCGAGSSSLERYATSNFDDTLNFIKTHPLMDSVPS:FNRPFLSTWYAYRL 420
421 1TKAVDTAAGPYQNTVVFILGSEKGIILKPLARLGNSGF:NDLFLFLEXSVYNSEKSYD 480
421 1TKAVDTAAGPYQNTVVFILGSEKGIILKPLARLGNSGF:NDLFLFLEXSVYNSEKSYD 480
481 1GVEKRTKMGXQLDRASSSLVAFSTCVIKVPLGRCHGKCKKTC:ASRDPYCGWIKEGG 540
481 1GVEKRTKMGXQLDRASSSLVAFSTCVIKVPLGRCHGKCKKTC:ASRDPYCGWIKEGG 540
541 1ACSHLSNSRLTFFQDIERGNTDGLGCHNSFVALNGHSSSLLPSTTSSDAGEYESR 600
541 1ACSHLSNSRLTFFQDIERGNTDGLGCHNSFVALNGHSSSLLPSTTSSDAGEYESR 600
584 1PSITTSSTAGEYESRGKGLDWGGLLSDPSTDP:GAVASHKHQKGVIRESY:KGHD 640
584 1PSITTSSTAGEYESRGKGLDWGGLLSDPSTDP:GAVASHKHQKGVIRESY:KGHD 640
601 1PSITTSSTAGEYESRGKGLDWGGLLSDPSTDP:GAVASHKHQKGVIRESY:KGHD 640
601 1PSITTSSTAGEYESRGKGLDWGGLLSDPSTDP:GAVASHKHQKGVIRESY:KGHD 640
644 1CLVPVTLATAVILAFVWGAVFSHTVYCHDHPKQVAVVQKKELEFHSRPGSSSYT 700
644 1CLVPVTLATAVILAFVWGAVFSHTVYCHDHPKQVAVVQKKELEFHSRPGSSSYT 700
661 1CLVPVTLATAVILAFVWGAVFSHTVYCHDHPKQVAVVQKKELEFHSRPGSSSYT 720
661 1CLVPVTLATAVILAFVWGAVFSHTVYCHDHPKQVAVVQKKELEFHSRPGSSSYT 720
704 1KLSGFGCTGSKDKKPBAILTFLMHKSKLATPNTAKX:IKAPCHHELTALTPSTPT 760
704 1KLSGFGCTGSKDKKPBAILTFLMHKSKLATPNTAKX:IKAPCHHELTALTPSTPT 760
721 1KLSGFGCTGSKDKKPBAILTFLMHKSKLATPNTAKX:IKAPCHHELTALTPSTPT 780
721 1KLSGFGCTGSKDKKPBAILTFLMHKSKLATPNTAKX:IKAPCHHELTALTPSTPT 780
764 1LOQRKPSRSGSREVERKQNTLNATKTPYNGSPVPTDLP:RAGSPSHIPSVVYVPTQ 820
764 1LOQRKPSRSGSREVERKQNTLNATKTPYNGSPVPTDLP:RAGSPSHIPSVVYVPTQ 820
781 1LOQRKPSRSGSREVERKQNTLNATKTPYNGSPVPTDLP:RAGSPSHIPSVVYVPTQ 840
781 1LOQRKPSRSGSREVERKQNTLNATKTPYNGSPVPTDLP:RAGSPSHIPSVVYVPTQ 840
824 1GYQHEVYDQPKYSEVAGMALDQATLEYKTIKEHLSKSPNGHVLVENDSLPDKVPQ 880
824 1GYQHEVYDQPKYSEVAGMALDQATLEYKTIKEHLSKSPNGHVLVENDSLPDKVPQ 880
841 1GYQHEVYDQPKYSEVAGMALDQATLEYKTIKEHLSKSPNGHVLVENDSLPDKVPQ 900
841 1GYQHEVYDQPKYSEVAGMALDQATLEYKTIKEHLSKSPNGHVLVENDSLPDKVPQ 900
884 1REASLGPTGASLSOTGLSKRLEMHSSSYGVYKESYPTNSL:PSHCATTKKNTNGSN 940
884 1REASLGPTGASLSOTGLSKRLEMHSSSYGVYKESYPTNSL:PSHCATTKKNTNGSN 940
901 1REASLGPTGASLSOTGLSKRLEMHSSSYGVYKESYPTNSL:PSHCATTKKNTNGSN 960
901 1REASLGPTGASLSOTGLSKRLEMHSSSYGVYKESYPTNSL:PSHCATTKKNTNGSN 960
944 1SSHLSRNQSGFRGNPPAPQAVRUSTVSHSSQSGQAVTVSRQPS:NAVNSLTSRGLKPT 1000
944 1SSHLSRNQSGFRGNPPAPQAVRUSTVSHSSQSGQAVTVSRQPS:NAVNSLTSRGLKPT 1000
961 1SSHLSRNQSGFRGNPPAPQAVRUSTVSHSSQSGQAVTVSRQPS:NAVNSLTSRGLKPT 1020
961 1SSHLSRNQSGFRGNPPAPQAVRUSTVSHSSQSGQAVTVSRQPS:NAVNSLTSRGLKPT 1020
1004 1PSLKFDVPPKPSFAPLSTSKKPNDACT 1040
1004 1PSLKFDVPPKPSFAPLSTSKKPNDACT 1040
1021 1PSLKFDVPPKPSFAPLSTSKKPNDACT 1047
1021 1PSLKFDVPPKPSFAPLSTSKKPNDACT 1047
```

## RESULT 2

S-09-957-187-4

Sequence 4, Application JS/0957187

Publication No. JS200300545141

GENERAL INFORMATION:

```
1 APPLICANT: Shinkels, Richard A.
2 APPLICANT: Laroche, William
3 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
4 FILE REFERENCE: 15366-540 CIP
5 CURRENT APPLICATION NUMBER: US/09/957,187
6 PRIOR FILING DATE: 2000-09-19
7 PRIOR APPLICATION NUMBER: 65/123,667
8 PRIOR FILING DATE: 1999-03-29
9 PRIOR APPLICATION NUMBER: 23/540,781
10 PRIOR FILING DATE: 2000-03-03
11 PRIOR APPLICATION NUMBER: 60/234,082
12 PRIOR FILING DATE: 2000-09-20
13 PRIOR APPLICATION NUMBER: 60/233,798
14 PRIOR FILING DATE: 2000-09-19
15 PRIOR APPLICATION NUMBER: 60/174,485
16 PRIOR FILING DATE: 2000-01-04
17 NUMBER OF SEQ ID NOS: 85
18 SOFTWARE: PatentIn Ver. 2.1
19 SEQ ID NO 4
20 LENGTH: 939
21 TYPE: PRT
22 ORGANISM: Homo sapiens
23 US-09-957-187-4
```

```
Query Match 90.34; Score 4921; DB 11; Length 939;
Best Local Similarity 95.64; Pred. No. C;
Matches 929; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MRSBALLLYPTLLHPAGAGPPEDSEP:SIHGNYTKQYFVGVGHKPGRNTTQRHRLDIOM 60
DB 1 MRSBALLLYPTLLHPAGAGPPEDSEP:SIHGNYTKQYFVGVGHKPGRNTTQRHRLDIOM 60
QY 61 1XIMNGILYIARPHITVYDICTSHTEETVCSKLTWKSQADVTCTRMKKGKDECHNF 120
DB 61 1XIMNGILYIARPHITVYDICTSHTEETVCSKLTWKSQADVTCTRMKKGKDECHNF 120
QY 121 1KVLKXGCEALFYCGTNAFNSFENYKMTLEFFDFSSVACQPYDAKHANVALFAQ 180
DB 121 1KVLKXGCEALFYCGTNAFNSFENYKMTLEFFDFSSVACQPYDAKHANVALFAQ 180
QY 181 1KYSATVTDFLAIDAVIYRS:GESPTLRTVKHDSKWKKEPYFQAVDGYIYFFREIA 240
DB 181 1KYSATVTDFLAIDAVIYRS:GESPTLRTVKHDSKWKKEPYFQAVDGYIYFFREIA 240
QY 241 1VEYNTMGKVVFPVAVQVCKXDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFYENILQAVT 300
DB 241 1VEYNTMGKVVFPVAVQVCKXDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFYENILQAVT 300
QY 301 1DVIRNGRDVVLATFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEKQSPSTWTFVPCR 360
DB 301 1DVIRNGRDVVLATFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEKQSPSTWTFVPCR 360
QY 361 1VPRPCCGAGSSSLERYATSNFDDTLNFIKTHPLMDSVPS:FNRPFLSTWYAYRL 420
DB 361 1VPRPCCGAGSSSLERYATSNFDDTLNFIKTHPLMDSVPS:FNRPFLSTWYAYRL 420
QY 421 1TKAVDTAAGPYQNTVVFILGSEKGIILKPLARLGNSGF:NDLFLFLEXSVYNSEKSYD 480
DB 421 1TKAVDTAAGPYQNTVVFILGSEKGIILKPLARLGNSGF:NDLFLFLEXSVYNSEKSYD 480
QY 481 1GVEKRTKMGXQLDRASSSLVAFSTCVIKVPLGRCHGKCKKTC:ASRDPYCGWIKEGG 540
DB 481 1GVEKRTKMGXQLDRASSSLVAFSTCVIKVPLGRCHGKCKKTC:ASRDPYCGWIKEGG 540
QY 541 1ACSHLSNSRLTFFQDIERGNTDGLGCHNSFVALNGHSSSLLPSTTSSDAGEYESR 600
DB 541 1ACSHLSNSRLTFFQDIERGNTDGLGCHNSFVALNGHSSSLLPSTTSSDAGEYESR 600
QY 601 1SGMLDNKHLSDPSTDP:GAVASHKHQKGVIRESY:KGHDQVPTLLAIVLAFV 660
DB 601 1SGMLDNKHLSDPSTDP:GAVASHKHQKGVIRESY:KGHDQVPTLLAIVLAFV 660
QY 661 1XGAVFSGITVYCVCHHARKDVAVVQKKELEFHSRSGSVSVTKLSGLFGDTQSKDKPKE 720
```

```

661 MGAVFSGITVVCVCHRRKQVAVIQKELTHSRGSMSSVTKLSGLFGDTQSKDKPE 720
721 AILTPVHNGKLAIPGNTAKMLKADQHLDLTAJTPSTPTLQCKKPSRGSRERN 780
722 AILTPVHNGKLAIPGNTAKMLKADQHLDLTAJTPSTPTLQCKKPSRGSRERN 780
781 QNLINACTKMPNGSFVPTDGLASASHIPSVVLPITQGYCHETVDQPKSEVAQ 840
782 QNLINACTKMPNGSFVPTDGLASASHIPSVVLPITQGYCHETVDQPKSEVAQ 840
841 MALEDQAATLEYKTIKERSSKSNHGNLVENLDSLPKVPQREASJGPPGASLSQTGL 900
842 MALEDQAATLEYKTIKERSSKSNHGNLVENLDSLPKVPQREASJGPPGASLSQTGL 900
901 SKLEMHSSSYGVYDKESYPTNSLTRSHQT 932
902 SKLEMHSSSYGVYDKESYPTNSLTRSHLT 977

RESULT 3
S-09-957-187-6
Sequence 6, Application US/09957187
Publication No. US20030054514A1
GENERAL INFORMATION:
APPLICANT: Shinkets, Richard A.
APPLICANT: Laroche, William
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT APPLICATION NUMBER: US/09/957,187
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/123,657
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/520,781
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/234,082
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 884
TYPE: PRT
ORGANISM: Homo sapiens
S-09-957-187-6

Query Match 84.2%, Score 4592.5, DB 11, Length 884;
Best Local Similarity 93.7%, Pred. No. 0;
Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;

1 MRSEALLVFTLHFAGAGPEDEPISISHGNVTKQVYFVGKPGRNTPQRR-LIQW 60
2 MRSEALLVFTLHFAGAGPEDEPISISHGNVTKQVYFVGKPGRNTPQRR-LIQW 60
3 MRSEALLVFTLHFAGAGPEDEPISISHGNVTKQVYFVGKPGRNTPQRR-LIQW 60
61 IMENGLYLAAPRYQNTVDTHTETVCSKLTWKSPQADVTCRMKGKIDCHNP 120
62 IMENGLYLAAPRYQNTVDTHTETVCSKLTWKSPQADVTCRMKGKIDCHNP 120
63 IMENGLYLAAPRYQNTVDTHTETVCSKLTWKSPQADVTCRMKGKIDCHNP 120
121 IKVLLKNDKDALFCVCGTNAFNSRNKMTLEPPGPEPSKARCPYDAKAVVA-PAQ 180
122 IKVLLKNDKDALFCVCGTNAFNSRNKMTLEPPGPEPSKARCPYDAKAVVA-PAQ 180
181 KLYSATVTDLAIDAVIYRSLGEPTLRTYKHSKWLKEPYFVQANDYGYVYFFREIA 240
182 KLYSATVTDLAIDAVIYRSLGEPTLRTYKHSKWLKEPYFVQANDYGYVYFFREIA 240
183 KLYSATVTDLAIDAVIYRSLGEPTLRTYKHSKWLKEPYFVQANDYGYVYFFREIA 240
241 VEYNTMKVVPFVAVQCKNDMGISQAVLQKQWTSFLKRLNCSYFGDSHFYNILCAVT 300
242 VEYNTMKVVPFVAVQCKNDMGISQAVLQKQWTSFLKRLNCSYFGDSHFYNILCAVT 300

```

```

301 DVIRINGRVYLATFSTPYNSIPGSAYCAYDMLDIAVFTGRKEKQSPGSTKTEPDER 360
302 DVIRINGRVYLATFSTPYNSIPGSAYCAYDMLDIAVFTGRKEKQSPGSTKTEPDER 360
361 VFKPRPGCCAGSSSSJERYATSNFPPDTLNFIKTHPLMDZAVESINRPMFLRTMVRYL 420
362 VFKPRPGCCAGSSSSJERYATSNFPPDTLNFIKTHPLMDZAVESINRPMFLRTMVRYL 420
421 TKIAYDAAEPYQNTVYFVLSSEKGIILKFLARIGNSGPLNDSLFLEMSVYNSKESYD 480
422 TKIAYDAAEPYQNTVYFVLSSEKGIILKFLARIGNSGPLNDSLFLEMSVYNSKESYD 480
481 GVEKRMGMOLDSSASSLYVAFSTCVIKPLGRCERHGKCKTCIASRFPYQWIKEGG 540
482 GVEKRMGMOLDSSASSLYVAFSTCVIKPLGRCERHGKCKTCIASRFPYQWIKEGG 540
541 ACSHLSNERSRTFFQDIERGNTDGLGCHNSFVALNGHSSLSLFTTSTSTAOEGYER 600
542 ACSHLSNERSRTFFQDIERGNTDGLGCHNSFVALNGHSSLSLFTTSTSTAOEGYER 600
543 ACSHLSNERSRTFFQDIERGNTDGLGCHNSFVALNGHSSLSLFTTSTSTAOEGYER 600
601 GGMLDKWKLDELSPSTDPLGSAVSSRHQDKGVRESYLKCHDCLVPTLLAIAVILAFV 660
602 GGMLDKWKLDELSPSTDPLGSAVSSRHQDKGVRESYLKCHDCLVPTLLAIAVILAFV 660
661 MGAVFSGITVVCVCHRRKQVAVIQKELTHSRGSMSSVTKLSGLFGDTQSKDKPE 720
662 MGAVFSGITVVCVCHRRKQVAVIQKELTHSRGSMSSVTKLSGLFGDTQSKDKPE 720
721 AILTPVHNGKLAIPGNTAKMLKADQHLDLTAJTPSTPTLQCKKPSRGSRERN 780
722 AILTPVHNGKLAIPGNTAKMLKADQHLDLTAJTPSTPTLQCKKPSRGSRERN 780
781 QNLINACTKMPNGSFVPTDGLASASHIPSVVLPITQGYCHETVDQPKSEVAQ 840
782 QNLINACTKMPNGSFVPTDGLASASHIPSVVLPITQGYCHETVDQPKSEVAQ 840
841 MALEDQAATLEYKTIKERSSKSNHGNLVENLDSLPKVPQREASJGPPGASLSQTGL 900
842 MALEDQAATLEYKTIKERSSKSNHGNLVENLDSLPKVPQREASJGPPGASLSQTGL 900
901 SKLEMHSSSYGVYDKESYPTNSLTRSHQT 932
902 SKLEMHSSSYGVYDKESYPTNSLTRSHLT 977

RESULT 4
US-09-957-187-30
Sequence 30, Application US/09957187
Publication No. US20030054514A1
GENERAL INFORMATION:
APPLICANT: Shinkets, Richard A.
APPLICANT: Laroche, William
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT APPLICATION NUMBER: US/09/957,187
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/123,657
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/520,781
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/234,082
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 30
LENGTH: 630
TYPE: PRT
ORGANISM: Homo sapiens
US-09-957-187-30

```

Query Match		61.7%	Score 3362	DB 11	Length 626
Best Local Similarity		100.0%	Prod. No. 5.2e-274		
Matches 626		Conservative	Mismatches 0	Indels 0	Gaps 0
US-09-957-187-63					
Query Match		61.7%	Score 3362	DB 11	Length 626
Best Local Similarity		100.0%	Prod. No. 5.2e-274		
Matches 626		Conservative	Mismatches 0	Indels 0	Gaps 0
US-09-957-187-63					
2Y	17	GAGFEDSEPISSHGNNTKQYVFGVGHKPGKPNNTORHRLDIOQIMMNGTLYIAARHIIT	78		
2Y	1	GSGFEDSEPISSHGNNTKQYVFGVGHKPGKPNNTORHRLDIOQIMMNGTLYIAARHIIT	60		
2Y	77	YVVDTSHTSEIYCSKLTWKSROADVTCRMKGKHQBCHNFIKVLKKNDDALFVCGTN	138		
2Y	61	YVVDTSHTSEIYCSKLTWKSROADVTCRMKGKHQBCHNFIKVLKKNDDALFVCGTN	120		
2Y	137	TNAFSPCRNYKVDLTLEPGDFSGMARCPYDAKHANVALFADGKLYSATVDFLAIDAV	198		
2Y	121	TNAFSPCRNYKVDLTLEPGDFSGMARCPYDAKHANVALFADGKLYSATVDFLAIDAV	180		
2Y	197	YVSGESPTLRVTKHDSKMLKEPYFVQADVTCRMKGKHQBCHNFIKVLKKNDDALFVCGTN	258		
2Y	181	YVSGESPTLRVTKHDSKMLKEPYFVQADVTCRMKGKHQBCHNFIKVLKKNDDALFVCGTN	240		
2Y	257	VCKNDGSGQSVLEKQWTSFLKARLNCSPGDSHFYFNILQAVTDVIRINGRDVVLATFS	318		
2Y	241	VCKNDGSGQSVLEKQWTSFLKARLNCSPGDSHFYFNILQAVTDVIRINGRDVVLATFS	300		
2Y	317	TPVNS:PGSAVCAYDMLDCLASVFTGRFKEQKSPDSTWTPDVPDVPKPRGCCAGSSLE	378		
2Y	301	TPVNS:PGSAVCAYDMLDCLASVFTGRFKEQKSPDSTWTPDVPDVPKPRGCCAGSSLE	360		
2Y	377	RVATNSFEPPDCLNFKTHPLMDEAVPS:FNRPWFLRTMVRVRLTKIAVDTAAGPYQNTTV	438		
2Y	361	RVATNSFEPPDCLNFKTHPLMDEAVPS:FNRPWFLRTMVRVRLTKIAVDTAAGPYQNTTV	420		
2Y	437	VVFLGSEKGIILKFLARIGNSGFLDLSLEBMSVYNSEKSYGVGDEKRNQVQJDRAS	498		
2Y	421	VVFLGSEKGIILKFLARIGNSGFLDLSLEBMSVYNSEKSYGVGDEKRNQVQJDRAS	480		
2Y	497	SSUYAFSTCVIKVPLGRCERHKGKCKT:CIASRDPCYGMKEGACSHLSPNSRLTFFEOIE	558		
2Y	481	SSUYAFSTCVIKVPLGRCERHKGKCKT:CIASRDPCYGMKEGACSHLSPNSRLTFFEOIE	540		
2Y	557	IBRGNTDGLGCHNSFVALNGHSSLSLPTSTTSCDSTAQGYESRGXLDKXJLSPDTP	618		
2Y	541	IBRGNTDGLGCHNSFVALNGHSSLSLPTSTTSCDSTAQGYESRGXLDKXJLSPDTP	600		
2Y	617	DLGAVSSRHHQCKKGVIRESYLKGHCQ	645		
2Y	601	DLGAVSSRHHQCKKGVIRESYLKGHCQ	626		
RESULT 5					
US-09-957-187-63					
Sequence 81, Application US/09957187					
Publication No. US2003054514A					
GENERAL INFORMATION:					
APPLICANT: Shimkels, Richard A.					
APPLICANT: LaRoche, William					
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY					
FILE REFERENCE: 5946-540 CIP					
CURRENT FILING DATE: 2000-09-19					
PRIOR FILING DATE: 1999-03-09					
PRIOR APPLICATION NUMBER: 09/520,781					
PRIOR FILING DATE: 2000-03-23					
PRIOR APPLICATION NUMBER: 09/234,092					
PRIOR FILING DATE: 2000-09-20					
PRIOR APPLICATION NUMBER: 60/233,795					
PRIOR FILING DATE: 2000-09-19					
PRIOR APPLICATION NUMBER: 60/174,435					
PRIOR FILING DATE: 2000-01-04					
NUMBER OF SEQ ID NOS: 95					
SOFTWARE: Patent in Ver. 2.1					
US-10-391-413-4					
Sequence 4, Application US/10391413					
Publication No. US20030167482A					
GENERAL INFORMATION:					
APPLICANT: Kimura, Toru et al.					
TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME					
FILE REFERENCE: 0020-5123P					
CURRENT APPLICATION NUMBER: US/0391,413					
CURRENT FILING DATE: 2003-03-19					
NUMBER OF SEQ ID NOS: 20					
SOFTWARE: Patent in version 3.1					
US-10-391-413-4					
Sequence 4, Application US/10391413					
Publication No. US20030167482A					
GENERAL INFORMATION:					
APPLICANT: Kimura, Toru et al.					
TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME					
FILE REFERENCE: 0020-5123P					
CURRENT APPLICATION NUMBER: US/0391,413					
CURRENT FILING DATE: 2003-03-19					
NUMBER OF SEQ ID NOS: 20					
SOFTWARE: Patent in version 3.1					
US-10-391-413-4					
Sequence 4, Application US/10391413					
Publication No. US20030167482A					
GENERAL INFORMATION:					
APPLICANT: Kimura, Toru et al.					
TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME					
FILE REFERENCE: 0020-5123P					
CURRENT APPLICATION NUMBER: US/0391,413					
CURRENT FILING DATE: 2003-03-19					
NUMBER OF SEQ ID NOS: 20					
SOFTWARE: Patent in version 3.1					
US-10-391-413-4					



```

PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644283
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816741
PRIOR FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/854203
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854290
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/899599
PRIOR FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: 09/928,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/23551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/50720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 35
LENGTH: 896
TYPE: PR
ORGANISM: Homo Sapien
US-09-931-836-35

Query Match 36.9%; Score 2009.5; DB 11; Length 896;
Best Local Similarity 45.2%; Pos: 7,4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 207

AY 5 ALLVFTLLHFAAGAFEDSEPISSHONYTKQYVFGKHKGENTTOR--HSLDQMY 62
b 12 ALLLLLLGGNNLFFPEEPPLSVARFDYNNYFVVGSGOPULTAREADUNLDRV 71
AY 63 IMNGLYAARDHRYVDIDTSHTSELYCSKLTWKPQADNDTCRKGKHKECHNFIK 122
b 72 RWNRTLPIDSDNLYRVLESEPTSTELRYQKLTWRSNPDINVCYKKGQKECHNFIK 131
AY 123 VLLKXNDALFVCSNAPNRCRYKVKDTLEPPGDSGMARCPYDAKHANVALPADGL 132
b 132 VLLRDESTLEFVCSNAPNRCRYKVKDTLEPPGDSGMARCPYDAKHANVALPADGL 131
AY 183 YSATVTDFLAIDAVYRSLGESPTLRTVGHDSKWLKEPYFVCAYVGDYIYFFPREIAE 242

```

192 PTAATVTOFLAIDAVYRSLGESPTLRTVGHDSKWLKEPYFVCAYVGDYIYFFPREIAE 251  
 QY YNTMGKVVPRVAQVCKNDXGGSORVLEKOWISFLKAPLNCSCVPDGSHPFNILCAVTDV 302  
 DS FNYLEKVVSRVAVCKNDXGGSPVLEKOWISFLKAPLNCSCVPDGSHPFNILCAVTDV 311  
 QY IIRNGROVVLATFTFYNSIPGSAVCAYOMLCLNSVFTGRFKEQKQFDSWTVPYDGRVP 362  
 DS VSLGGRPVYLVAFVFTESNSIPGSAVCAYOMLCLNSVFTGRFKEQKQFDSWTVPYDGRVP 371  
 QY KPRGCCAGSSSLERFATNSNFFDGLNFKTHPLMDHVAVPSIFMRKEDLSTVYRYSLETK 422  
 DS RPRGCCAAGPM--CYNASSALPDGLNFKTHPLMDHVAVPSIFMRKEDLSTVYRYSLETK 429  
 QY IAVDTAAGFYQNHVTFVSGSEKGIILKFLAR--LONSGLFNLDS--FLEMSVYNSEKSYD 480  
 DS VAVDVAGAGPMQNVVTFVSGSEKGIILKFLAR--LONSGLFNLDS--FLEMSVYNSEKSYD 489  
 QY GVED--KEIYQMQLDRASSSLYVAFSTCVIKVPLGRGERHKGKCKTCLASRDPIYCGWINE 538  
 DS GGGETGQGLLSLELCAASGGLLAAFFRCVVRVPVAVRCQYSGCYKNGICIGSQDPYCGWAPD 549  
 QY GGACSHLSNSRLTFEQDIBRGNTDGLGCHNSFVALNGHSSSLPSTTTSDSTAQBOYE 598  
 DS GSCIFLSFGTRAAFEQDVSGASTSGLODC----- 578  
 QY SRGMYLWKHLSDSPDSTPLGAYSSHHNQDKKGVIRESVLKGHDQLVPVTLIAIVILA 658  
 DS GLLRASLSEDRAGLVSVNLLVTSVAA 606  
 QY FVNGAVPSSITV-YQVCDHPRKQVAVVCRKEKE--LTHSRGSSSVSVTKLSGLFGDQSK 715  
 DS FVNGAVPSSITV-YQVCDHPRKQVAVVCRKEKE--LTHSRGSSSVSVTKLSGLFGDQSK 715  
 QY KPEAILTPVHNGKCATPGNTAK-MLIKADQHLQJLTALPTPEST 761  
 DS WAKATLLOGSPHLDLSGLPTPEQT 713  
 QY PTLQQRKP-----SGSRWERNQNLINACTKMPMGSPVITDLP---LRASPSH 811  
 DS PTLQQRKP-----SGSRWERNQNLINACTKMPMGSPVITDLP---LRASPSH 811  
 QY IPSVVVLPITQGGYCHYVQPKMSEVAQXALEQQA--LEVKTIKEHLSSKSPHNGVLY 871  
 DS APEQPPAPGEPPTDGRUYAARPGHSGHSGOFFLTTHASPDREVV-----SAPTGPJCPA 911  
 QY ENLDSIP-PKVPQREAF-----GPGASLSQT 896  
 DS SAADGLPRENSPPTGSLRRTLGPHAPPAALTRAT 846

RESULT 8  
 US-10-035-977-35  
 : Sequence 35, Application US/10035977  
 : Publication No. US20030134327A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Desnoyers, Luc  
 : APPLICANT: Batton, Dan L.  
 : APPLICANT: Goddard, Audrey  
 : APPLICANT: Godowski, Paul J.  
 : APPLICANT: Gurney, Austin L.  
 : APPLICANT: Pan, James  
 : APPLICANT: Stewart, Timothy A.  
 : APPLICANT: Watanabe, Colin K.  
 : APPLICANT: Wood, William I.  
 : APPLICANT: Zhang, Zemin  
 : TITLE OF INVENTION: SECRETED AND TRANSMBRANE POLYPEPTIDES AND NUCLEIC  
 : TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 : FILE REFERENCE: P303081C1C  
 : CURRENT APPLICATION NUMBER: US/10/035,977  
 : CURRENT FILING DATE: 2001-12-26  
 : PRIOR APPLICATION NUMBER: 60/065579



PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/112514  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113300  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/113430  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113605  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113624  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/114140  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/115558  
PRIOR FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/116043  
PRIOR FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/125774  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/125778  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/125936  
PRIOR FILING DATE: 1999-03-24  
PRIOR APPLICATION NUMBER: 60/127035  
PRIOR FILING DATE: 1999-03-31  
PRIOR APPLICATION NUMBER: 60/127706  
PRIOR FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: 60/129122  
PRIOR FILING DATE: 1999-04-13  
PRIOR APPLICATION NUMBER: 60/130359  
PRIOR FILING DATE: 1999-04-21  
PRIOR APPLICATION NUMBER: 60/131270  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/131272  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/131294  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/132371  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/132379  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/132383  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138166  
PRIOR FILING DATE: 1999-06-06  
PRIOR APPLICATION NUMBER: 60/144794  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/146970  
PRIOR FILING DATE: 1999-08-03  
PRIOR APPLICATION NUMBER: 60/162506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 09/311832  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 09/386142  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 09/644848  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 09/747259  
PRIOR FILING DATE: 2000-12-20  
PRIOR APPLICATION NUMBER: 09/816744  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: 09/854208  
PRIOR FILING DATE: 2001-05-16  
PRIOR APPLICATION NUMBER: 09/854280  
PRIOR FILING DATE: 2001-05-10  
PRIOR APPLICATION NUMBER: 09/874503  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: 09/865539  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: 09/9081827  
PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: PCT/US99/10733  
PRIOR FILING DATE: 1998-05-14  
PRIOR APPLICATION NUMBER: PCT/US99/28551  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30720  
PRIOR FILING DATE: 1999-12-22  
PRIOR APPLICATION NUMBER: PCT/US99/05601  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: PCT/US99/05841  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: PCT/US99/14042  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: PCT/US99/15264  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/US99/23522  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: PCT/US99/23329  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: PCT/US99/32678  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: PCT/US99/34956  
PRIOR FILING DATE: 2000-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/06520  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: PCT/US99/17800  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: PCT/US99/19692  
PRIOR FILING DATE: 2001-06-23  
PRIOR APPLICATION NUMBER: PCT/US99/21066  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: PCT/US99/21735  
PRIOR FILING DATE: 2001-07-09  
NUMBER OF SEQ ID NOS: 50  
SEQ ID NO 35  
LENGTH: 888  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-035-977-35

Query Match 36.94; Score 2009.5; DB 12; Length 888;

Best Local Similarity 45.24; Pred. No. 7.4e-160;

Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY	5	ALLLYFTLLHEAGAGFEDSEPISSHCNVTIKQYVNVGHKPGHNTTOR--HRLDIQMIM 62
DB	12	ALLGELLJGGARSFPSEEPPLSVAPROYLNHYFVPVVGSGPGLTPAEGADJLNIQRLV 71
QY	63	IMNGTIVIAARDHIYTVTIDTSHTEEVCSKKJWKSRQADVDTCRMKGKHKECHNFIK 122
DB	72	RVNRTLFIGDRDNLRYVELAEPTSTELRYCHKLFRSNPSPDINVCRMKGKGESECRNFVK 131
QY	123	VLLKKNDDALFVCGTNAFNFSCRNKXMTLEPPGDEFGSMARCPYDAKHANVALFADGKL 182
DB	132	VLLLRDESTLFVCGSNAPFPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALFSDGML 191
QY	193	YSATVTOFLATDAVIYSSLGESPTLRVKHDSKULKEPYFVCAVDYGYIYFFREIATVE 242
DB	132	FTATVTDFEA:DAVIYRSLGDRPTLRVKHDSKAFKBPYPVHAVESHSYVFFREIATME 251
QY	243	YNTGKVVFPRVAQVCKNMNGSGSORVLEKQWTSFLKARLNCSPVGDGSHFYFNTCAVTDV 302
DB	252	FNYLEKVVSVAVARVCNDCVGSFVLEKQWTSFLKARLNCSPVGDGSHFYFNVLCVATGV 311
QY	303	IRINGRQVLAFTSTPVNS:PGSAVCAYDMEDIASVTTGRFKESKSPDSTWTPVDFDRVP 362
DB	312	VSLGGRPVLAFFSTPSNS:PGSAVCAPDLTQVAAVFEGREFREKSPESIWTPVPEQVP 371
QY	353	KPRGCCAGSSSLEFYATSNFEPDCTLNFKTHLMDENAVPSIENRPNFLRMVRYRLTK 422
DB	372	RPRGCCAARQNM--CYNASSALPDIDLNFKVTHLMDENAVPSLHAPMILATLKHQLTR 429
QY	423	IAVETAAAGPYCNHTVVF:GSEKGIILKFLAR--IGNSFNLNDSLFLEMSVYNSEKSYD 480

```

b 430 VANDVGGKPKGKCTVWPLGSSAGVILKEJUNPNASTSGTISGLSVFLIBETVTPDEURP 489
y 481 GVED--KRIMQQLDRASSVYVAFSTCV:KVPDGRCHRGKQKCTIASRPFYQOWKE 538
b 490 GGSETGGLLSLE:DAASGGLLAAPFCVVPVVARCOOYGGCKVXCTSGOBYGMAHD 548
y 539 GGACSHLSPNSRPTFFGCIERGTNDLGDCHNSVALNHSGLLSTTTERSTAGQGE 599
b 550 -GSCIFUSPOTRAAFECQVSGASTSGDC----- 678
y 599 SPCGMLEKHLSDSDSTOM:GAVSSHNOCKAGVIFESYLKRGDQ:NPVLLAIIVILA 658
b 579 -----TG:LASLSEBRAG:VSVNLLTSSVAA 606
y 659 FVNGAVPSGTV-YVCHHRKQVAVVQRKEK:--LTHSRGMSVTKLSGLFGDTQSK 715
b 607 FVNGAVPSGTVGAVFVGLRERELA--PRKDKEA:LAHGAGEAVLSVRL:--GERRAQ 660
y 716 DP-----KPEAILTELVYNGK:LATPONTAK-M:IKADQHHLDTALPTTEST 761
b 661 GPGRGGGGGGAGVPTFALLAPLMQNG-----WAKATLQGGPHLDSGLLPTFEQT 713
y 762 PTLOQKRP-----SRGSRWERNQNLINACTKMPFVSGSPVITLFP-----PRASPSH 811
b 714 P-LPKR:PTPHPHAL:GPEAWCH-----GHPLLPASASSSLLAPAR 757
y 812 PSVVVJ:PTQGGYCHEVVDQPKYSEVAQVALESQAATLEKTK:KEHLSKSPHGMV 871
b 758 APEQPPAFGEPTFEGRYAARFGRASHGDCFLTTHASFDHRRVY-----SAPTSELDPA 811
y 772 ENLDSLP-PRVPOREAS:-----GPGASLSOT 898
b 812 SAADGLPRWSPPTGSLRFLGPHAFPAITLRT 846

```

## RESULT 9

S-10-137-870-544

Sequence 544, Application US/10137870

Publication No. US2003015893A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen.

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

APPLICANT: DeBoer, Laura

```

APPLICANT: Deforge, Laura
APPLICANT: Deeneyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C158
CURRENT APPLICATION NUMBER: US/10-140,021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
S-10-140-021-544
Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7,4e-160;
Matches 423; Conservative 130; Mismatches 242; Indels 141; Gaps 20;
Y 5 ALLAYTLHLFAGAGPDEPSTISHGNVTKQYVFGVGHKPGRNTTQR--HRLDIQIMV 62
b 12 ALLALLLLGGAGS:FPBEPPLSVAPROY.NHVPVFGSGPGLTPAEGADDLNIQRLV 71
Y 63 IMNGTLTYAARDHNYTVYDICTSHTEEYCSKLTWKGRQADVTCRMKGKHKDECHNPIK 122
b 72 RVNRTLFIQDRKLYRVELEPPTSTELRYQKLTWRGNSPDINVCRMKGKGEGRNPFVK 131
Y 123 VLLKKNDALFVCGTNAFNSCRNYKMTLEPFGDEFSGVARYPYDAKHAVALFADGKL 182
b 132 VLLRCESTLFGVGSNAFNPVCANYSIDTQYGVNLSGMARCFYDEKHAVALFADGML 191
Y 193 YSATVDFDAIDAVIVRSUGESTLRTVYHDSKWKKEFYFVQAVQYGGYIYFFPREIAYE 242
b 192 FTAIVTDFDAIDAVIVRSUGESTLRTVYHDSKWKKEFYFVQAVQYGGYIYFFPREIAYE 251
Y 243 YNTMGVVFPRVAQCKNDMGSGORVLEKQWTSPLKARLNCSPGSPSHFVENILQAVTV 302
b 252 FRIEXVTVSVRVARCKNDVGGSRVLEKQWTSPLKARLNCSPGSPSHFVENILQAVTV 311
Y 303 I-R-NGRDVVLATEPYNISPGSAVCAYDVLDIASVFTQPFKEQKSPDSTWTPVPERVE 362
b 312 VSLGGRPVAVESTESNIPSGSAVCAFDLTQVAAVFGRFRFQKSPESWTPVEDQVE 371
Y 363 KPRFGCCAGSSSLERYATSNFPDPTLNATKTPLMDEAVPSIPRPNFPLRTVRYRLTK 422
b 372 RPRFGCCAGPXY--QYNASSAEPDILNPKTPPLXDEAVPSLSGHAPMTJRTLMRHLTR 429
Y 423 IAYDTAAGPYONHTVYVFLGSEKGIILKFLAR--IGNSGFNDLSLPLEMSVYNSEKSYD 430
b 430 VAVDVGVGPNQGVTVPLGSEAGTVLKLFLVRYNASTSGSGSVFEEFETRYRDKGAP 449
Y 481 GVED--KRIMGMOLDRAESSLYVAFSTGVIKVPLGCRERHKKATCIASRPVCGNINE 538
b 490 GGGETGCRLLSLJELDAASGELLAAFPYCVVYVAVRQQVQSGCMKICIGSDPYCGNAPF 549
Y 539 GGACSHLSPNSRVTFFODTERGNTDGLGCHNSFVALNGHSSSLAPSTTTSSTAGEVE 598
b 550 -GSCIFLSPQTRAAFEODVSGASTSGSLGDC----- 578
Y 599 SRGGMLEWKHLLESPDSTDFPLGAVSSHNDCKKGVRESYVKGHDQVVPVTLIAIVILA 658

```

```

DB 579 -----TGLLRASLSFDRAGLVSNLLTSSVAA 606
QY 659 FWGAVFSGITV-YVCDHRRKQVAVVQRKEKE--LTHSRRGSMSSVTKLSGLFQDQSK 715
DB 607 FWGAVVSGFSVGMFVGLRERRELA--RRKQKEAILAHGAGEAVLSVSRLL-----GERRAQ 660
QY 716 DP-----KPEAILTPVHNKGLATPGNTAK-XLIKADQHHLDJIALFTPEST 761
DB 661 GPGRGGGGGGAGVPPPEALAPLVQNG-----WAKATLLQGGPHDLOGSJLFTPEQT 713
QY 742 PTLQCKRKE-----SRGSRWERNQNLINACTKQKPMGSPVPTDLP---LRASPSH 811
DB 714 P-LQCKRKEFTPEPHALGSRANDH-----CHPJJFASASSSJLLAPAR 757
QY 812 IPSVVVLPITQGYQYHEYDQPKMSVAQMALEDAATLEYKTKKEHLSSKSPNHNVLV 871
DB 758 APEOFFPAPGEPTDGRLYAARPGASHGDPPELTPEHASEDFRRVV-----SAP*GLDPA 811
QY 872 ENLSLP-PKYQPREASU-----GPGASLSQT 898
DB 812 SAADGLPEPWSPPFTGSLRRLGPHAPPAATCRRT 946

```

```

RESULT 1:
US-10-140-021-544
Sequence 544, Application: US/10-140,021
Publication No. US20030138886A:
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Deeneyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C167
CURRENT APPLICATION NUMBER: US/10-140,021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-021-544

```

```

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7,4e-160;
Matches 423; Conservative 130; Mismatches 242; Indels 141; Gaps 20;
QY 5 ALLAYTLHLFAGAGPDEPSTISHGNVTKQYVFGVGHKPGRNTTQR--HRLDIQIMV 62
DB 12 ALLALLLLGGAGS:FPBEPPLSVAPROY.NHVPVFGSGPGLTPAEGADDLNIQRLV 71
QY 63 IMNGTLTYAARDHNYTVYDICTSHTEEYCSKLTWKGRQADVTCRMKGKHKDECHNPIK 122
DB 72 RVNRTLFIQDRKLYRVELEPPTSTELRYQKLTWRGNSPDINVCRMKGKGEGRNPFVK 131
QY 123 VLLKKNDALFVCGTNAFNSCRNYKMTLEPFGDEFSGVARYPYDAKHAVALFADGKL 182

```

132 VLLURDESTLVGSGRAFNPCVAVSYDITLCPYGENISVAVKCPYQPKANVAFDQEM 131  
 183 YSATVDFDLADAVIRSLGSPRLTVYKHSKWKKEPFPYQVANDGQDNYFFPRRIAE 142  
 192 FTATVDFDLADAVIRSLGSPRLTVYKHSKWKKEPFPYQVANDGQDNYFFPRRIAE 151  
 243 YNMGKVVPRVAVQCKNMGSPVLEKQWTSFLKAKLNSVPGSHSVENILQAVTGW 102  
 252 FNYLEKVVSRVAVQCKNMGSPVLEKQWTSFLKAKLNSVPGSHSVENILQAVTGW 111  
 303 IRNGRNVVATSTPNSIPGSAVQAVYLDIASVFTPFKPKSKSDSTWTPVDERVP 162  
 312 VSLGGRVWLVAVSTPNSIPGSAVQAVYLDIASVFTPFKPKSKSDSTWTPVDERVP 171  
 363 KPRPGCCAGSSLSERYATSNFEDTLNFVKTUPLHDEAVPSLEKWPWILATVVRILTK 122  
 372 RPRPGCCAGSSLSERYATSNFEDTLNFVKTUPLHDEAVPSLEKWPWILATVVRILTK 131  
 423 IAVDTAAGPYQNHVTFVFGGSEKGTILKFLAP--IGNSGPLVNSLFLDEKSVYNSKSCSYD 182  
 430 VAVDVAGPWNQVTFVLSSEAGTVLKFVPRNAS--SOTSGLSVFLSEETVPRDCCGRP 191  
 481 QVED--KAIKMGQJDRASSSVYVAFSTCVIKVPLGRGERHKGKKTCTIASRDPYCGMIKE 242  
 490 GGBETGQSLLSLELDAASGGGLAAFFRCVVRVAVCOQYSGCKNCLGSDPYCGWAFD 251  
 539 GGACSHLSPNSRLTFEQDIERGNTDGLGDCHNSFVALNGHSSLSLPTTTTSCSTACQSYE 302  
 550 -GSCIFLSPGTRAAAFECQVSGASTSGJGDC----- 311  
 599 SRGGMLDKWHLLDSEFSTDPGLVYSHNHQDKKGVIRESVYKGHDCQVPTLLAIAVILA 362  
 579 -----TGLRASLSSEDRAGLVSVNLLVTSSVAA 606  
 659 FVMGAVFSGITV-YVCQDHRKRVAVVQRKEKE--JTHSRRGSSSVTKLSGLFGDQOSK 715  
 607 FVMGAVFSGITV-YVCQDHRKRVAVVQRKEKE--JTHSRRGSSSVTKLSGLFGDQOSK 660  
 716 DP-----KPEALITPLMHNGKATPONTAK-XLIKADQHHLDTALPTPEST 761  
 661 GPGRGGGGGGAGVPEALPLKMGQ-----WAKATLQGGPHDLSGLLPTPEQT 713

## RESULT 12

US-10-140-274-544

Sequence 544, Application US-10-140-274

Publication No. US20030143674A1

## GENERAL INFORMATION:

Applicant: Baker, Kevin P.

Applicant: Beresini, Maureen

Applicant: DeForge, Laura

Applicant: Desnoyers, Luc

Applicant: Filvaroff, Ellen

Applicant: Gao, Wei-Qiang

Applicant: Gerritsen, Mary E.

Applicant: Goddard, Audrey

Applicant: Godowski, Paul

Applicant: Gurney, Austin

Applicant: Sherwood, Steven

Applicant: Smith, Victoria

Applicant: Stewart, Timothy A.

1 : APPLICANT: Tomas Daniel  
 2 : APPLICANT: Kananabe, Colin K  
 3 : APPLICANT: Wood, William  
 4 : APPLICANT: Zhang, Zelin  
 5 : TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 6 : FILE REFERENCE: P3330R1C161  
 7 : CURRENT APPLICATION NUMBER: US/10/140,274  
 8 : PRIORITY DATE: 2002-05-06  
 9 : PRIORITY APPLICATION: See file wrapper of P3330R1C161  
 10 : NUMBER OF SEQ ID NOS: 550  
 11 : SEQ ID NO 544  
 12 : LENGTH: 838  
 13 : TYPE: PRT  
 14 : ORGANISM: Homo Sapien  
 15 : US-10-140-274-544

Query Match 36.9%, Score 2009.5, DB 12, Length 838

Best Local Similarity 45.2%, Pred. No. 7.4e-162

Matches 423, Conservative 130, Mismatches 241, Indels 141, Gaps 20

QY 5 ALLVFTLLHFAGAGFPEDSEPSISHQNYTKQYVYVGHKPGNTTCR--HRLDICIMM 62  
 DB 12 ALLLLLLLGAGAGLFPBEPPLSVAFRDYLNHYVFPVSGSGRLTPAEGADDLNIQVRL 71  
 QY 63 IYNGTLVYIAARCHIYTVDDIDTSHTEBYCSKLLTKSRQADVTCRMYKSKKDECHNPIK 122  
 DB 72 RVNRTLFIGDRNDYRVELEPTSTELRYQKLLWRNPSDINCYRMKGQEGECRNPKV 131  
 QY 123 VLLKNDALVCGTNAFNPSCRNKVKQDTLEPPFDFSGMARCPYDAKHANVALFADGKL 182  
 DB 132 VLLURDESTLVGSGRAFNPCVAVSYDITLCPYGENISVAVKCPYQPKANVAFDQEM 191  
 QY 183 YSATVDFDLADAVIRSLGSPRLTVYKHSKWKKEPFPYQVANDGQDNYFFPRRIAE 242  
 DB 192 FTATVDFDLADAVIRSLGSPRLTVYKHSKWKKEPFPYQVANDGQDNYFFPRRIAE 251  
 QY 243 YNMGKVVPRVAVQCKNMGSPVLEKQWTSFLKAKLNSVPGSHSVENILQAVTGW 302  
 DB 252 FNYLEKVVSRVAVQCKNMGSPVLEKQWTSFLKAKLNSVPGSHSVENILQAVTGW 311  
 QY 303 IRNGRNVVATSTPNSIPGSAVQAVYLDIASVFTPFKPKSKSDSTWTPVDERVP 362  
 DB 312 VSLGGRVWLVAVSTPNSIPGSAVQAVYLDIASVFTPFKPKSKSDSTWTPVDERVP 371  
 QY 363 KPRPGCCAGSSLSERYATSNFEDTLNFVKTUPLHDEAVPSLEKWPWILATVVRILTK 422  
 DB 372 RPRPGCCAGSSLSERYATSNFEDTLNFVKTUPLHDEAVPSLEKWPWILATVVRILTK 429  
 QY 423 IAVDTAAGPYQNHVTFVFGGSEKGTILKFLAP--IGNSGPLVNSLFLDEKSVYNSKSCSYD 480  
 DB 430 VAVDVAGPWNQVTFVLSSEAGTVLKFVPRNAS--SOTSGLSVFLSEETVPRDCCGRP 489  
 QY 481 QVED--KAIKMGQJDRASSSVYVAFSTCVIKVPLGRGERHKGKKTCTIASRDPYCGMIKE 538  
 DB 490 GGBETGQSLLSLELDAASGGGLAAFFRCVVRVAVCOQYSGCKNCLGSDPYCGWAFD 549  
 QY 539 GGACSHLSPNSRLTFEQDIERGNTDGLGDCHNSFVALNGHSSLSLPTTTTSCSTACQSYE 599  
 DB 550 -GSCIFLSPGTRAAAFECQVSGASTSGJGDC----- 578  
 QY 599 SRGGMLDKWHLLDSEFSTDPGLVYSHNHQDKKGVIRESVYKGHDCQVPTLLAIAVILA 658  
 DB 579 -----TGLRASLSSEDRAGLVSVNLLVTSSVAA 606  
 QY 659 FVMGAVFSGITV-YVCQDHRKRVAVVQRKEKE--JTHSRRGSSSVTKLSGLFGDQOSK 715  
 DB 607 FVMGAVFSGITV-YVCQDHRKRVAVVQRKEKE--JTHSRRGSSSVTKLSGLFGDQOSK 660  
 QY 716 DP-----KPEALITPLMHNGKATPONTAK-XLIKADQHHLDTALPTPEST 761  
 DB 661 GPGRGGGGGGAGVPEALPLKMGQ-----WAKATLQGGPHDLSGLLPTPEQT 713

```

y 762 PTJQCKRKP-----SRSGREWRKCNLINACTKMP22MGSSVPTDLP-----LAASRSH 811
o 714 P-LPCKRKPTEHPHHAALGERAWH-----GHFLPASASSSLLJAPAR 757
y 812 PSVVVLRITQGYQYQYVQCPKNSSEVAQMALEDOATLEYKTIKEHLSSKSPNKGVLV 871
b 758 APEQPPAPGEPFDGRLVAARQFASGSPPLTPHASPERRVW-----SAPTGLDPA 811
v 872 ENLDSLP-PKVQREASJ-----GPFQASLSCT 998
b 812 SAADGLPREWSPPTGSLRRPLGHPAPPAATLRRT 846

RESULT 13
S-10-140-471-544
; Sequence 544, Application US/10/140.87
; Publication No. US20030134354A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3338163
CURRENT APPLICATION NUMBER: US/10/140.471
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper of Patm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
S-10-140-471-544

```

```

Query Match 36.9%; Score 2099.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred No. 7.4e+160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

y 5 ALLYPTLHFAAGAPFEPSEPTISGSHKNTKQYVPVGHKPKRNTTQ--HRLTIQMLY 62
b 12 ALLLLLLGGAGLPEBPPLSVAPRYLNNYPVVGSGGR-LTPAEGADLQVRL 71
y 63 INGTLYLAARDHYIVDITDTSHTBIYSKLLTWKSRQADYDTQRMKKHKBDSHPFK 122
b 72 RVNRTLFICGRNLYRVELEPTSTJRYCRKLTWRSNEDINVCRMKQKQEGCRNFVK 131
y 123 VLLKKNDQALFVCGTWAFKSPCNRYKMDLTPEPFGDFSGNARCPYDAKHANVALPADGKL 182
b 132 VLLRDESLFVCGSNAPFVCANYSIDTLOPGVGN--SGMARCPYDPKXANVALSDGML 191
y 183 YSNATVDFLAIDAVIKYS--GESPTLTVKHDSKWLKREPYFVCAVQYGVYVFFPREIAVE 242
b 192 FTATVDFLAIDAVIYSLGRDPTLRVXHSKAWKEPYFVHAVGSHVYVFFPREIAVE 251
y 243 YNTYKGVYFPRVACVKNMGSGQRVLEKQWTSFLKARLNCQSVGESHFYENILQAVITY 302
b 252 FNYLERVVAVRVACVKNMGSGQRVLEKQWTSFLKARLNCQSVGESHFYENILQAVITY 311
y 303 IRNGREVLVAFSTPYNS:PGSNVANDYLDIASVFTQHPHKKSPDSTNTFVPCDEAVF 362

```

```

DB 312 VSLOGRPVLAFTSPNSIPGSAVCAFDLTQVAAVEGFRFBOKSPESWTVPEDQVP 371
QY 363 KPRPGCAGASSSERYATSNFEFDDTLNFIKTHPLMDEAVFSINRPMFLRTMVRYELTK 422
DB 372 RPRPGCAAQPM--QYNASSALPDEILNFVKTHPLMDEAVESLGHAPWILRTLMPHQLTR 429
QY 423 IAVDTAAGPQNHTWVFJGSEKGIILKFLAR--IGNSGFLNDSLFJBEEMSVYNSEKSYD 480
DB 430 YAVDVGAGPWNQTVWFLGSEAGTVLKLVRPNASTSGTSLGSVJFEETFYRFDRCGRP 489
QY 481 GVED--KRIMGMOULDRASSGLYVAFSTCVIKVPLGRCERHGKCKKTCIASRDPYCGMIKE 538
DB 430 GGETSORLISJELDAJSCGLLAAPFPCVVRVVARCOQYSGCKMKNCGSDPYCGWAPC 543
QY 539 GGACSHJSPNSRJTFEODIFRGNTGLGDCCHNSFVALKNGHSSSLPFTTTSDSTAQSGYE 598
DB 550 -GSCIFLSPGTRAAFEQOVSGASTSGLGDC----- 578
QY 599 SRGGMLDKKHLLDSPDSTDFLGAVSSHNDKQKVIRESYLKGHDQLVPTLLAIAVILA 658
DB 579 -----TGLLRASLSEDRAGLVSNVLLVTSVAA 606
QY 659 FWGAVFSGITVAYCVCDHRRKQVAVYQKXEKE--LTHSRRGSMSSVTKLSGLFGDTQSK 715
DB 607 FVGVAVVSGFSGVGFVGRERRELA--RRKDKEALLAHGAGEAVLSVRL---GERRAQ 660
QY 716 DP-----KFEAILTPXHNKJLATPONTAK-KLIKADQHLLDLTAJFTTEST 761
DB 561 GPGGFGGGGGGAGVPRPEALLAPLVQNG-----WAKATLLOGGPHCLDSGLLFTPEOT 713
QY 762 PTJQCKRKP-----SRSGREWRKCNLINACTKMP22MGSSVPTDLP-----LPASPSH 811
DB 714 P-SPKELPTPHPHHAALGERAWH-----CHPLLEFASASSSLLJAPAR 757
QY 812 IPSVVVLRITQGYQYQYVQCPKNSSEVAQMALEDOATLEYKTIKEHLSSKSPNKGVLV 871
DB 758 APEQPPAPGEPFDGRLVAARQFASGSPPLTPHASPERRVW-----SAPTGLDPA 811
QY 872 ENLDSLP-PKVQREASJ-----GPFQASLSCT 998
DB 812 SAADGLPREWSPPTGSLRRPLGHPAPPAATLRRT 846

```

```

RESULT 14
US-10-140-807-544
; Sequence 544, Application US/10/40807
; Publication No. US20030134354A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3338163
CURRENT APPLICATION NUMBER: US/10/140.807
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See File Wrapper of Patm
NUMBER OF SEQ ID NOS: 550

```

[illegible]

```

RESULTS
US-10-140-922-544
: Sequence 544, Application US/10140922
: Publication No. US200303889A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Destroyers, LLC
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerltzen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurrey, Austin J.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watarabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zelin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P333031C179
: CURRENT APPLICATION NUMBER: US/10/140,922
: CURRENT FILING DATE: 2002-05-07
: Prior Application removed - See P.a.m or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 544
: LENGTH: 858
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-140-922-544

```

Query Match	36.9%	Score	2009.5	DB 12	Length	682			
Best Local Similarity	45.2%	Pred.	No. 7.4e-160						
Matches	423	Conservative	130	Mismatches	241	Indels	141	Gaps	20
Qy	5	ALLVFTLLHPRAGAPFEDSPISISGNYTKYVPVGHKPGKNITQR---HRLD-QMIM 62							
Db	12	ACLLLLGSLGSHGUFPEPPFSLVAPRDVLNHYVFEVSGPGLRLPAEGADDJNTQYL 71							
Qy	63	INKGLVYLAARCHIVTQVDTSH-EETVCSKKLTKKRGQADVTCCKMKGKIKDCHECNFIK 122							
Db	72	RNRTJLFGDRDLNVALELEPFTSELYRQKLTWRNPSDINVCKMKSKCEGECNFKV 131							
Qy	123	VLLKKNDALFVGTNATNPSCRNYKMDLTLEPGEDEFSGMARCPYDAKHANVALFADGKL 182							
Db	132	VLLRDESTLFCGSKAENPVCANVSIDLQPVGDNISGVARCPDPFKANVALFSGMG 191							
Qy	183	YGATVTDLDAIVYRSLGSEPTLRVCKHDSKWLKEPYEQVQAVDVGCDVLYFFFEIAYE 242							
Db	192	FPAQVTDLDAIVYRSLGSEPTLRVCKHDSKWLKEPYFVHAVEMGSHVYFFFEIAYE 251							
Qy	243	YNTGKAWFPRVACVCKNDMGSSGRVLEKQWTSFLKARLNCSVPGDGHFVNILQAVTV 302							
Db	252	PVYLEAVTVSRVARVCKNDVSGSFVLEKQWTSFLKARLNCSVPGDGHFVNVLQAVTV 311							
Qy	303	IPNGRCVVLATFTSPVNSIPGSACVAVMDLIASVPTGRFCKQKSPDSTWTPVDPERV 362							
Db	312	VSLGGEPTVLAVFSTPNSIPGSACVAFDLTQVLAHVPEGRFCKQKSPESLWTPVPEQVP 371							
Qy	363	KRPGCCAGSSSLEAYATSNFPDCTLNFVTKRPLXDEAVPS-FNRPWF-RTMVPVYRLTK 422							
Db	372	KRPGCCAPGK-QYNASSALPEDILNFVKTHPLXDEAVPSLGHAPWILRTLUMRQHLR 429							
Qy	423	IAVUTAAEPYQNHVYVFGSGSKG--LKLPLAR--IGNSGFLNCSLFLLEBMSVYNSEKCSVD 480							
Db	430	VAVDVGAGFMWCTVYVFGSSAGATVLFVLVRPNASTIGTSLGSLVF-LEETETVYRPRDCRGP 489							



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer i.i.d.  
OM protein - Protein search, using sw mode:  
Run on: October 23, 2003, 17:09:07 ; Search time 19 Seconds  
(without alignments)  
2293.694 Million cell updates/sec  
US-09-856-681-2  
Perfect score: 5450  
Sequence: 1 MRSEALLFTLLHFAGASPF.....PPKFSFAPLSTXKXNDACT 1030  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 328717 seqs, 42310953 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Issued Patents AA:  
1: /cgn2\_6/prodata/2/iaa/5A.COM5.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5S.COM5.pep.\*  
3: /cgn2\_6/prodata/2/iaa/5A.COM3.pep.\*  
4: /cgn2\_6/prodata/2/iaa/5S.COM3.pep.\*  
5: /cgn2\_6/prodata/2/iaa/5A.COM5.pep.\*  
6: /cgn2\_6/prodata/2/iaa/5S.COM5.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2013.5	36.9	888	US-09-077-940A-4	Sequence 4, Appl
2	1997	36.6	887	US-09-077-940A-2	Sequence 2, Appl
3	1519.5	27.9	930	US-09-254-594-6	Sequence 6, Appl
4	1428.5	26.2	929	US-09-254-594-3	Sequence 3, Appl
5	969.5	17.8	730	US-08-121-713D-58	Sequence 58, Appl
6	969.5	17.8	730	US-08-835-268-59	Sequence 58, Appl
7	969.5	17.8	730	US-09-060-610-58	Sequence 58, Appl
8	969.5	17.8	730	US-08-833-391-58	Sequence 58, Appl
9	969.5	17.8	730	US-09-060-610-58	Sequence 58, Appl
10	969.5	17.8	730	PCT-US94-10151A-58	Sequence 58, Appl
11	880	16.1	650	US-08-121-713D-60	Sequence 60, Appl
12	880	16.1	650	US-08-835-268-60	Sequence 60, Appl
13	880	16.1	650	US-09-060-610-60	Sequence 60, Appl
14	880	16.1	650	US-08-833-391-60	Sequence 60, Appl
15	880	16.1	650	US-09-060-610-60	Sequence 60, Appl
16	880	16.1	650	PCT-US94-10151A-60	Sequence 60, Appl
17	852.5	15.6	655	US-08-835-268-54	Sequence 54, Appl
18	852.5	15.6	771	US-08-121-713D-54	Sequence 54, Appl
19	852.5	15.6	771	US-09-060-610-54	Sequence 54, Appl
20	852.5	15.6	771	US-08-833-391-54	Sequence 54, Appl
21	852.5	15.6	771	US-09-060-610-54	Sequence 54, Appl
22	852.5	15.6	771	PCT-US94-10151A-54	Sequence 54, Appl
23	852.5	15.6	771	US-08-121-713D-54	Sequence 54, Appl
24	845	15.5	712	US-08-835-268-64	Sequence 64, Appl
25	845	15.5	712	US-09-060-610-64	Sequence 64, Appl
26	845	15.5	712	US-08-833-391-64	Sequence 64, Appl
27	845	15.5	712	US-09-060-610-64	Sequence 64, Appl

26	845	15.5	712	1	US-09-060-610-64	Sequence 64, Appl
29	845	15.5	712	5	PCT-US94-10151A-64	Sequence 64, Appl
30	779.5	14.3	724	1	US-08-121-713D-62	Sequence 62, Appl
31	779.5	14.3	724	1	US-08-835-268-62	Sequence 62, Appl
32	779.5	14.3	724	2	US-09-060-610-62	Sequence 62, Appl
33	779.5	14.3	724	3	US-08-833-391-62	Sequence 62, Appl
34	779.5	14.3	724	4	US-09-060-610-62	Sequence 62, Appl
35	779.5	14.3	724	5	PCT-US94-10151A-62	Sequence 62, Appl
36	748	13.7	775	4	US-09-308-179B-1	Sequence 2, Appl
37	736.5	13.5	862	4	US-08-556-422A-2	Sequence 2, Appl
38	654.5	12.0	607	4	US-08-556-422A-4	Sequence 4, Appl
39	586	10.8	477	3	US-08-136-922-2	Sequence 2, Appl
40	539	9.9	425	4	US-08-556-422A-7	Sequence 7, Appl
41	462	8.5	295	4	US-08-556-422A-6	Sequence 6, Appl
42	437	8.0	634	3	US-09-041-236-2	Sequence 2, Appl
43	437	8.0	634	4	US-09-771-467C-2	Sequence 2, Appl
44	437	8.0	666	3	US-09-240-410-2	Sequence 2, Appl
45	356.5	6.5	606	3	US-09-041-236-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-09-077-940A-4  
Sequence 4, Application US/09077940A  
Patent No. 6576441  
GENERAL INFORMATION:  
APPLICANT: KIMURA, Toru et al.  
TITLE OF INVENTION: NOVEL SEAPHORIN 2 AND GENE ENCODING THE SAME  
FILE REFERENCE: 0010-4426P  
CURRENT APPLICATION NUMBER: US/09/077,940A  
CURRENT FILING DATE: 1998-06-25  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: Patent version 3.1  
SEQ ID NO 4  
LENGTH: 889  
TYPE: PPT  
ORGANISM: Homo sapiens  
US-09-077-940A-4  
Query Match 36.9%, Score 2013.5; DB 4; Length 989;  
Best Local Similarity 45.3%; Pred. No. 1.7e-182;  
Matches 424; Conservative 129; Mismatches 241; Indels 141; Gaps 20;

QY	5	ALLLYFLHFGAGFEDSEPI	ISHGNTKQVPVGHKPGRNTTOR--HRJDIQIM	62
DB	12	ALLLLLLGGAGGJFEDPP	PLSVAPRDYLNHYVFGSGPGRLT	PAEGADCLNIQRVL 71
QY	63	IMNGTLVIAARDHIYTV	DIOTSHTEIYCCKLTWKSROADVOTCRMKGHKDCHEKFIK	122
DB	72	RVNRTLFGIDRNLVY	VELEPPSTELAYORKLTWRNSPDSINVCVRKKGQEGECRNVK	131
QY	123	VLLKNNIDALFVCGTNA	PNFNSCRNYKMDTLEPFDEFGSMARCPYDAKANKVALFADGKL	182
DB	132	VLLLRDSESTLFCVGSNA	PNVCANYSIDTLPQVGDNISGVARCPYDPKHANVALFSDGML	191
QY	183	YSATVTDLADAVIYSL	SGESPTLRTVKDLSKWLKEPYFVCAVDYGYIFPEREIAVE	242
DB	192	FTATVTDFLADAVIYSL	SGDRPTLRTVKHDSKWKKEPYFVHAEVWGHSHYFFREIAVE	251
QY	243	YNTGKVVFEVACVCKND	YGGSGGVLEKQWTSFLKARLNCSPVGDSPGHFVNLCQAVTDV	302
DB	252	PNYLEKVVSVARVCND	YGGSGSPVLEKQWTSFLKARLNCSPVGDSPGHFVNLCQAVTG	311
QY	353	IRINGRDVLTATFT	SPYNS:PGSAVCAIDMLD:ASVFTGRFKQKSPDSTWTVPVDSRVP	362
DB	312	VSLGRPVLLAVFST	PSNS:PGSAVCAFDLTQVAAVFESGRFREQKSPESIWTPVFEQVP	371
QY	363	KPRGCCAGSSSLRY	ATSNEFFDDTLNFKTHTPLVDEAVFSPNRPFLRTMVPYRLTK	422
DB	372	RPRGCCAAPGV--	QYNASSALPDDILNFVKHTPLVDEAVFESGHAPKILRTLNRHQLTR	429



```
QY 425 IAVDTAGPQNHVTFVGGSEKGIILKFLAR--IGNSGFLNDSLPLEKSYVNSEKSYD 480
D6 430 VAVDVGAGPKNQWFLGSEAGNVLKPLVRRNAS--SGTSLGLVFLERETTFYFPCGCRP 489
QY 481 GVFD--KRIIMQQLDRASSSLVAFSTCVKVPUGRGERHKKKCIASRPPQYQWKE 538
D6 490 GGGTGGRIILSLERFASGGLLAAPFRCVVPVPCVQYSGGNKNCISGQYQWAPD 549
QY 539 GGACSHSPNHRUTPEDIERTGNTDGLGCHNSFVALNGHSSSLJSTTTSDTAGRGEY 596
D6 550 -GSCIFLSPGTAAAFEDVSGASTSLGDC----- 578
QY 599 SRGCVLKKHLDSPOGTDPLGAVSSHHOKKQVRESYLKHTQLVPTLLAAVAILA 558
D6 579 -----TGLLPASLSDRAGLVSYNCLVTS 604
QY 659 FVMAVRSGLTVYVCDHRRKQVAVYORKEKE--LTHSRRGSMSSVTKL 705
D6 607 FVMAVRSGLTVYVCDHRRKQVAVYORKEKE--LTHSRRGSMSSVTKL 705
QY 716 DP-----KPEALTEPMXNKLATPNTAKMLIKADCHHLDLTALPTPEST 763
D6 661 GPGGRGGGGGAGVPEALLPLVQNG-----WAKATLLOGPHDIDSGLLPTPECT 715
QY 762 PTLCKEKP-----SRGSEWERNQNLNACTK-----DMPKXGSPVZPTD 802
D6 714 P-LPCKSLPTTHPHALGRFRHHA-----PADSTHL---LPGTSEPTAPPVP 887
QY 912 PPSVVVLPITQGYOYEVDOQKXSEVACMALEDCAATLBYKTKHEHSSKSPNHGWL 871
D6 758 APECPAPAGEPTPGRLVYARIGASHGDFPLTHASPDPRVV-----SATGGLDPA 811
QY 872 ENLSLP-FKVPQREASL-----GPGASLSGT 856
D6 912 SAASGLPRKSPPTTGLRRLRPLGHPAPPAATLRRT 849

RESULT 2
US-09-077-940A-2
; Sequence 2, Application US/09077940A
; Patent No. 6576441
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
; FILE REFERENCE: 2020-44269
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 887
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-077-940A-2

Query Match 35.6%, Score 1997, DB 4, Length 887;
Best Local Similarity 42.7%, Filed No. 63e-1a1;
Matches 433; Conservative 132; Mismatches 280; Intels 120; Gaps 25;

QY 2 PSEALLSYETLIIFAGAGPPELSPISIGNYTKQYFVYGHKRGNTTQF--HRLDIQ 59
D6 10 PPALLFILLLRVTHGILFPDEPPLSVAPRYSLSHYEVGSGPGRCTPAERAGDILITQ 69
QY 60 MINIMNGTLYAARDHVTVDICTSHRETVCSKMLTKSRCAVDCTCRKGHKDECHN 119
D6 70 RVLVRNRTFLGDRNCLQVSELESTSTELAYQRLKLRWSNPSSDLYCRKKGCGECRN 129
QY 120 FKVLKKNDCALFVCGTNAPNPGCRKYNKMTJEFDFEFGVAPCPYDAKHANVALFAD 179
D6 130 FVKVLLRDESTLFCVCGSNAPNPIKANSMDTLLGDNISGVARCPYDPKHANVALFSD 189
QY 190 GKLYSATVDFLDAIVLYRSLGSEPTLRTYKHSKWLKPEYFVQAVDYGDYTYFFEREI 239
```

```
D6 190 GMLFTATVTDFLAIDANLYRSLGDRPTLRTVKHDSKWKPEPYFVHAVWGSHVYFFEREI 249
QY 240 AVFYNTGKVFEPVAVGVCKNDXGGSORVLEKQWTSFLKARLNCSSVFGDSHFVFNILQAV 299
D6 250 AMEPNYLEKVVSVSVARVCKNDXGGSORVLEKQWTSFLKARLNCSSVFGDSHFVFNILQAV 309
QY 300 TCVIRIENGRDVIATFTSTPYKSIIPGSVAVCAVDMLDIASVETGFPECKSPDS--WTPVPDE 359
D6 310 TQVSLGGRTVILAVFTSTPNSIPGSVAVCAPDMQVAAVFEGRPREKSPES--WTPVPDE 369
QY 360 RVKPPRECCAGSSSLERYATSNFPDCTLNFIKTHPMDIEAVPSIENRDPWFJRTYVYR 419
D6 370 QVPRPRGCCAARGH--QYNASNALPDEILNFVKTHELMQENAVPSLGHSPVJATL--RHO 427
QY 420 TKTAVIDTAAGFYQYNTVTFJGSEKGIILKFLAR--IGNSGFLNDSLPLEKSYVNSEK 477
D6 428 LTRVAVDVGAGPKNQWFLGSEAGNVLKPLVRRNAS--SGTSLGLVFLERETTFYFPCGCRP 487
QY 478 ---SYDGYEDKRIKMXCLDRASSSLVAFSTCVKVPUGRGERHKKKCIASRDPYCG 534
D6 488 GRSSSAGEWGQRIILSLERFASGGLLAAPFRCVVPVPCVQYSGGNKNCISGQYQWAPD 547
QY 535 WKEGSGACSHSPNHRUTPEDIERTGNTDGLGCHNSFVALNGHSSSLJSTTTSDTAGR 594
D6 548 WAPD-GSCIF-LRPGTSATFEQVSGASTSLGDC----- 580
QY 595 EGYESRGMLEKHLDSPOGTDPLGAVSSHHOKKQVRESYLKHTQLVPTLLAAVAILA 654
D6 581 -----TGLLPASLSDRAGLVSYNCLVTS 604
QY 655 VTLAFVMAVRSGLTVYVCDHRRKQVAVYORKEKE--LTHSRRGSMSSVTKL 705
D6 605 SVAPFVMAVRSGLTVYVCDHRRKQVAVYORKEKE--LTHSRRGSMSSVTKL 705
QY 706 --SGJFGDTQSKDPKPEALTEPMXNKLATPNTAKMLIKADCHHLDLTALPTPEST 763
D6 663 GTGGRGGGGGAGVPEALLPLVQNG-----WAKATLLOGPHDIDSGLLPTPECT 715
QY 764 LCKEKP-----SRGSEWERNQNLNACTK-----DMPKXGSPVZPTD 802
D6 716 -LPCKSLPTTHPHALGRFRHHA-----PADSTHL---LPGTSEPTAPPVP 887
QY 803 -----LPLRASPSPHPVAVVDFITQGYOYEVDOQKXSEVACMALEDCAATLBYKTK 856
D6 721 PESRLCARRSCASHGDFPLTHASPDPRVV-----SATGGLDPA 811
QY 857 EHLSSKSPNHGWLNLSDLP-FKVPQREASL-----GPGASLSGTGLSKLEKHS 909
D6 817 -----DCLFGMSPFPATSSDRPGRPHGPTAALRRRT 849
QY 910 SSGYDYKSYFTNSLTPSHOATLTKRNTNNSKSSHLNRNCSFGRCD-NPPPAF 963
D6 850 FNSG-----EAPFGGHRFRHHA-----PADSTHL---LPGTSEPTAPPVP 887

PSEULT 3
US-09-054-594-6
; Sequence 6, Application US/09054594
; Patent No. 6566784
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 2020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Tissue Type: Child Brain
; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Identification Method: p for resulting peptide
JS-09-254-594-6

Query Match      27.0%  Score 1519.5; DB 4; Length 930;
Best Local Similarity 35.9%; Pred. No. 24e-135;
Matches 391; Conservative 135; Mismatches 156; Indels 189; Gaps 32;

QY 6 LLAYFTLLHPAGAGPEDSEPLISGNYTKQVPEVGHKPGKNTTORHRLDIOYVIMN 65
DB 13 LLLLSLSPH-TQAAFCQDPLISLQTSPLSWPFGLEDRAVAHEL-GLDQRFILN 70
QY 66 GTLYIAARDHIYVICDTSHT-BEIVCSKLTWKSRQADVTCRMKKHKECHNFYKVL 124
DB 71 RTLLVAARDHVFSFDLQAEDEEGSLVENKLTWRSQ--DMENCAVRGKLTDECYNVIRVL 127
QY 125 LKXNDALFVCGTNAIPSCRNKNDLEPFDERSKARCYDAKHANVALFADGKLYS 184
DB 129 VPMDSQTLACGNSFPCVRSYGITSLQGEELSGQARCPFDATOSTVAISAEGLYS 187
QY 185 ATVTDFLAIDAVIYRSLGESPTLRITVKHDSKWLKEPVYQAVDYGVYIFFFREIAVEYN 244
DB 189 ATADFQASDAVYRSLGPOPPLES AKYDSKMLREPHFVVALEHGDHVYFFLEKSL-WR 246
QY 245 T--MGKVPRPVAVQCKVMDGSGORVLEKQWTFJKAALNCSVPGDSHFYFNILQAVDV 302
DB 247 TPGUGRVQFSRAVAVCKRMGSGSPRADJDRHWTSPFKLRLNCSVPGDSTFYFDVLSQSTGP 306
QY 303 IRINGRDVLTATSTPNYSIPGSAVCAVMDLTASVFTGRFKSKQSPDSTWTPVDERVP 362
DB 307 VNLGRSALFGVFTTQNSIPGSAVCAFYDDCIERGEGFKSKQSLDGAWTPVSEDKVP 366
QY 363 KPRGCCAGAGSSSLERVAETNEFPDDTJNF-KTHPMDAEAVPSIFNRWPF-RTMVYRLTK 422
DB 367 SPRGSCAGVGAALPSSSQDLDDVLLFIKAPLLDPAVPPATHQP-JJTLISRALETC 425
QY 423 IAVTAAGPYONHTVFLGSEKGI::LKFLARINGSGLNDSLFLEEMSVYNSEKCVGV 482
DB 426 VAVDMAGPHRNTVLFEGSNDGTVLKVLPP-GQSLGSPFIILEEDDAYSHARCS--GK 482

RESULT 4
US-09-254-594-3
; Sequence 3; Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEXAPHORIN GENE; SEXAPHORIN Y
; FILE REFERENCE: 0030-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Tissue Type: Brain
; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Identification Method: p for resulting peptide
US-09-254-594-3

Query Match      26.2%  Score 1428.5; DB 4; Length 929;
Best Local Similarity 34.0%; Pred. No. 11e-126;
Matches 360; Conservative 147; Mismatches 355; Indels 197; Gaps 32;

QY 6 LLAYFTLLHPAGAGPEDSEPLISGNYTKQVPEVGHKPGKNTTORHRLDIOYVIMN 65
DB 13 LLLLSLSPH-TQAAFCQDPLISLQTSPLSWPFGLEDRAVAHEL-GLDQRFILN 69
QY 66 GTLYIAARDHIYVICDTSHT-BEIVCSKLTWKSRQADVTCRMKKHKECHNFYKVL 124
DB 70 RTLLVAARDHVFSFDLQAEDEEGSLVENKLTWRSQ--DMENCAVRGKLTDECYNVIRVL 127
QY 125 LKXNDALFVCGTNAIPSCRNKNDLEPFDERSKARCYDAKHANVALFADGKLYS 184
DB 128 VPMDSQTLACGNSFPCVRSYGITSLQGEELSGQARCPFDATOSTVAISAEGLYS 187
QY 185 ATVTDFLAIDAVIYRSLGESPTLRITVKHDSKWLKEPVYQAVDYGVYIFFFREIAVEYN 244
DB 189 ATADFQASDAVYRSLGPOPPLES AKYDSKMLREPHFVVALEHGDHVYFFLEKSL-WR 246
QY 245 T--MGKVPRPVAVQCKVMDGSGORVLEKQWTFJKAALNCSVPGDSHFYFNILQAVDV 302
DB 247 TPGUGRVQFSRAVAVCKRMGSGSPRADJDRHWTSPFKLRLNCSVPGDSTFYFDVLSQSTGP 306
QY 303 IRINGRDVLTATSTPNYSIPGSAVCAVMDLTASVFTGRFKSKQSPDSTWTPVDERVP 362
DB 307 VNLGRSALFGVFTTQNSIPGSAVCAFYDDCIERGEGFKSKQSLDGAWTPVSEDKVP 366
QY 363 KPRGCCAGAGSSSLERVAETNEFPDDTJNF-KTHPMDAEAVPSIFNRWPF-RTMVYRLTK 422
DB 367 SPRGSCAGVGAALPSSSQDLDDVLLFIKAPLLDPAVPPATHQP-JJTLISRALETC 425
QY 423 IAVTAAGPYONHTVFLGSEKGI::LKFLARINGSGLNDSLFLEEMSVYNSEKCVGV 482
DB 426 VAVDMAGPHRNTVLFEGSNDGTVLKVLPP-GQSLGSPFIILEEDDAYSHARCS--GK 482
```

```

483 ED--KRINGQJQDRASSLYVAFSTCVIKVPLGRCERKCKKTCIASRDPYQCHKE 338
483 RSRAARRIIGLEETGCHLPLFAFPGCIYLSLRCARHGACORSLASLDYQCHRRF 542
539 GGACSH:SPASPAATFEODIE-KGNTDZL--GCGHNSFVALKGRSSLSLSTTSTDAQR 595
543 RGVNIRGPGS-----TDVHILGCEKHEKD-QCG-----ATGQSQSPGSS 544
596 GYSEGCXLDWKHLDSQDSTPLGAVSSHHQKGVFEESTLKHQDQVPTLLAIAY 555
598 AY-----GVSECLSPASGASRSPFILLAY 610
656 ILAFVMAVESGTTVWCDEHKEKAVAVORKEKE-TSFRSGMSSVTKLSGLFGTQSK 115
611 AAFALGASVSLVSCAC--FRAN-----RRNSKLETPGLRPLSLRLARJHGGSEFP 664
716 DKRP---EAILTFMANGKLAIPONTAKML-KADQHELDLTALPTPSTPTLQCKEKR 722
665 PPPKDCAACTPOLYTFPLFPEGSGPP-----PLACLPSTTPELPVKULRAS 715
773 GSEKERNQNLINACTKNDPMVSPVPTDLP---LRASFSHIPSVVVPLTQQ---GYQ 926
716 GG-FWEMWONGNASEGPRGPRSCSAAGSPAPRAVVRPPPPGCGQGEVEVITLLEELRYL 724
927 H-----EYVDQEKYSEVAYALSDQAATLEYKTKHLSKSKENHGNNLV 871
775 HGQPPKSGSEPLASAPFTSPASPGALFVD-----SSNPR-----814
972 ENLDSUPP-----KVFQREASLPGPGASLSQTLGSKLEMHSSSYGVYKRY 920
815 ---DCVPEPLRVPPDGKRAAPSGRPALSAPAPRLGVSG-SRRL-----PF 856
921 PTNSLRTSHQATLTKRNTNTHSSNSLSRNSQSGRGNPPAPQVDS---IQVHSQPSG 979
857 PT-----HRA-----PPGLTRYSPGGRYSGQGR 863
979 QAVTVSRQPSLNAVNSLRSGLKRTPSLKPDV--PPKPS 1015
934 HLLYLGR-PGCHGRSLKAVDVKSPISPKPLATPQFA 921

RESULT 5
US-08-121-713D-58
; Sequence 58, Application US/08121713D
; Patent No. 563856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthews, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphore Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3100
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121.713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 5.4
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627

```

```

; REFERENCE/DOCKET NUMBER: B94-302-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4342
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-121-713D-58

Query Match 17.8%, Score 369.5, DB 1; Length 730;
Best Local Similarity 34.4%, Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 117; Gaps 29;

QY 5 LVTLLHFAAGAFPEDSEFIS:SHGNVTQKYPVFVGHKPGKNTTCR-----HSLDQMI 61
DB 11 LNWALH-AAANWDVSP-----KMVQF-----GEERVQFJGNESHKDHFKLL 54
QY 62 NIMGQTLVIAAPDHIYTVQID--TSHTTEIYCSKLTWKSRQADVDTCRMKGRKDECHN 119
DB 55 EKCHNSLLVGARNIVYANISLPDLTFTEQ-----RISWHSSGAHRELCTYLGKXSDDCQN 109
QY 120 FIKVLLKKNDDALFVCGTNAFNPSCRNYKMDTLEPFGD-----SPSGMARCPYDAKANV 174
DB 110 YIRVLAKIDDDRVLCGTNAYKPLCRHYALKD---GDYVVEKEYEGRGCLPFPDCHNST 165
QY 175 ALFADGKLYSATVDFLAIDAVIYRSLGESPTLRTVKHSKWLKEPYFVQAVDYGVYIF 234
DB 166 AYSEGQLYSATVADFSTDELIVRG-----PJTERSDLKQLNAPNFVNTMEYNDIFPP 220
QY 235 FFRZAEVYNTMGVFRVACVQKNDYGGQSRVLEKQWTSFLKARLNCVSPGDSHFYFN 294
DB 221 FFERATAVEYVCGKATYSRVARVCBKDGPHQGGDR-W-SFLXSRLNCVSPGDPYFVN 279
QY 295 ILQAVTVIRN-N-GRDY---VLATESPYNSIPGSAVCAYDMLDIASVFTGRPEQKSPD 350
DB 290 EQSTSDILEGNVGGQVEKLIYGVFTIPVNSIGSAVCAFSMKSLSEFDGPFKEQETYN 339
QY 351 STWTPVDPDERVPKPRPGCCAGSSSDERYATSNRFPDQTLNFIKTHPLMGZAVPPINRPW 410
DB 340 SNWLAVPSLKVPSPRQOCVND-----SRCLPQVSNVAFVKSHLMDZAVZAFFTFE1 391
QY 411 FLRTVRYRLTKIAYD---TAAGPYQNHVTYVFLGSEKGILKFLARIGNSGFYN----- 461
DB 392 LIR-SLOYRFTKIAYDQVKTPDG--KAYDVLFITDQGVKIKAL-----NSASFSDDTV 445
QY 462 DSLFLEMSVYNSEKSGVGDVEDKRIYGMQ-DRASSSLYVAFSTCV:KVPILGR- ERHG 519
DB 446 DSVI:BELOVLP-----PGVFWXLYVVRMGDDSKLVVYSDDELIAIKJHRCGSKIT 499
QY 520 KCKTCTIASHRDPYCNWIKERGACSHL-SPN---SRJTFEQDIERGNTDGLGDCNHFVAL 575
DB 500 NCRE-CVSLQDPYCAWKNVELKCTAVGSPKMSAGKRRFIQNSLGEHKA-CGGRQTEIV- 557
QY 576 NCHSSSLPSTTTSESTA-----QSGYSESGKMLDKWHLJDSQDSTPLGAVSSHHQ 628
DB 558 ---ASPVPTQPTTKSSGSDPMIS:HOAEFEPE---DNEIIVGVDDSNVPIPTLAEINRA 610
QY 629 DKKGVIRESVYKGHQDQVVPY---TJ-LAI-----AVI:AFVNGAVFS 666
DB 611 GSK-----LPSQEKLPYTAETJTAIVTSCLGAVVGVFISGFLFS 652

RESULT 6
US-08-835-268-58
; Sequence 58, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.

```

APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/635,268  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Conor, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 730 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-856-268-58

Query Match 17.8%; Score 969.5; DB 1; Length 730;  
Best Local Similarity 34.4%; Pred. No. 4.5e-83;  
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 23;  
QY 8 LVFTLLHFAGAGFPEPDSSEISHSNVTQYVFGVGHKPGKRNITQR-----HRLDQIM 61  
DB 11 LKVALH--AAAVNDVSP-----RMVGF-----GSEPVRFLGNESHKORFKLS 54  
QY 62 MINMGTVIAARDHIVTVID--TSHTTEIYCSKLTWKSROADVTCRMKGKKGDECN 119  
DB 55 ERDHSLSLVGARNIVNISRLDTERTEQ-----RIENHSSGARELCYLKSKSDDCN 109  
QY 120 FIKVLKXDDA-FVCGTNAFNSCRNYKMTLEPGD-----EFGXARCYPYDAKHAHV 174  
DB 110 VIRVLAKIDDDRLICOTRAYKLCRHVALKD---GVYVEKEVEGRGLCFDCHNST 165  
QY 175 ALPADCKLVSATVTDLAIDAVIYRS:GSPST:RTVYHDSK:KBPYFQAVCYGYIYF 234  
DB 166 AIYSEGLYSATVADFSG:DLPIYRG-----PLRTESDLQ:NAFNFNTKEYNDF:FF 220  
QY 235 FFEZ:AVENTMGSVVFPRAVACVKNDVGSORVLEKOWTSP:KARLNC:SVPGDSHFYFN 294  
DB 221 FRETAVEINCKAIYSARVAVCKKCGKGGPHOGGR-KTSP:KSLNCSVPGDFFYFN 279  
QY 295 ILCAVTDVIRIN-GRGV---VLAFTSTPYNS:FGSAVCAYDKL:DIASVFTGRFKQKSPD 350  
DB 280 E:QSTSDIIEGNVGGQVEKLIYVFTIPVNS:IGSAVCAFSKKS:ILESFGGPFKEQETWY 339  
QY 351 STWTPDVERVAPRGCCAGSSSJERYATSNFFPDITLNF:KTP:KDEAVFSP:PNRW 410  
DB 340 SNMLAVFS:KVPSPRGQCVCND-----SRTLDPVSVYFKSHT:KXDEAVFAFTRI 391  
QY 411 FLRTMVRRLTKIADV-----CANGPYCNHTVFLGSEKGIIL:KFLAR:IGNSGFLN----- 461

DB 392 LIRISLQYRRTKIADVCOVR:PDG--KAYDVLEGTDDGKVIKAL-----NSASFSSSTV 445  
QY 462 DSLF-BEMSVYNSEKSYGVVEDKXRMGMOLDRASSSLYVAFSTCVIKVPLGRG--ERHG 519  
DB 446 DSVV:EEBQLVLP-----PGVPVKVLYVMRMCGDSKLVVVSDEILAI:KLHRCGSKIT 499  
QY 520 KCKYTCIASRDPYCGK:KEGACASHL:SPN---SRLTFEQD:ERGNTDGS:GDCHNSEVAL 575  
DB 500 KCRE-CVSLQDPYCAWNVKXCTAVGSPDMSAGRRFIONISLGEHKACGGRPQTIV- 557  
QY 576 NGHSSLLPSTTSDSTA-----QEGYESRGGYLDWKHLLDSDPDSTDP:LGAVSSHNHQ 628  
DB 558 ---ASPVTQPTTKSGGDPVHSIHQAEFEPE---IDNEIVIGVCDNVP:ENTLAEINHA 610  
QY 623 EKKGVIRESV:KGHQDLVPY---TL-LAI-----AVLAFVYMGAVFS 666  
DB 611 GSK-----LPSSEKLPYVTAETLTIAIVTSCLGALVVGFS:SGFLFS 652  
RESULT 7  
US-09-060-692-58  
Sequence 58, Application US/09060692  
Patent No. 5935565  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Koidekin, Alex L.  
APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/360,692  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Conor, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 730 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-692-58  
Query Match 17.8%; Score 969.5; DB 2; Length 730;  
Best Local Similarity 34.4%; Pred. No. 4.5e-83;  
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;  
QY 6 LVFTLLHFAGAGFPEPDSSEISHSNVTQYVFGVGHKPGKRNITQR-----HRLDQIM 61

```
DB 11 LKVALH--AAANQNVSP-----KMYVQF-----QSERVQRPFGNESHKQHFLL 54
QY 62 MINNGTLVIAARSHIYVTD--TSHTTEIYCSKKLTWKERQADVTCRMKGKHKDECHN 119
DB 55 EKDNHSLLVGARNVYVNISRLUTEFEQ-----RIEWHSSGARRELOVYKKGSEDDCON 103
QY 120 FIKVLLKNDALFVCGTNAFNSCRNYKMDTEPFC-----EFSQWACRCPYDAKHANV 174
DB 110 YIRVLAKIDDERVLCGTNAVYKPLCRHVALKD---GDVVVEKEVEGRLGCLPDPDCHNST 165
QY 175 ALPADGKLYSATVDFLAIDAVIYRSLGESPTLTUVKHSKWLKEPYVQAVDGYIYF 234
DB 166 AYYSEGQYSATVADFSGTDFLIYRG-----PLRTSRDLKQLNAPNFVNTMEYNDPIFF 220
QY 235 FFRFIAVEYNTMGVPRVAQVCKNDMGSGQRVLEKQWTSFLKARLNCVPGDSHYFN 294
DB 221 FFRFIAVEYNTMGKAIYSRVARVCKDKGPHQGGDR-WTSFLKSRJLNCVPGDYPFFN 279
QY 295 ILQAVTVDIRIN-GRDY---VLATFTSPYNSIPGSAVCAVDYDVLASVFTGRFKEQKSPD 350
DB 290 BIQSTSDITBGNVGGQVEKLIYGVFTTPVNSIGSSAVCAFSMKSLIESFDGPFKEQETMN 339
QY 341 STWTFVPRDERVPRKRCGCCAGSSSLERYATSNBEPDCLNFIKTHPLMDRAVPSIRNPW 410
DB 340 SKWLAVPSCKVPEPRGQCVND-----SRTLPDVSVNPFVKSHTLMDRAVPAFFTRPI 391
QY 411 FLRTWVRVRLTKIAVD---TAAGFYQNHVTVFVGSSEKGIILKFLARIGNSGFJN----- 461
DB 332 LIR-SLQVRFTKAVQCVQVATPDG--KAYDVLFIQTGGKVIKAL-----NSASFDSDTV 445
QY 462 DSJFLEMSVYNSEKSYDGVDERKRMNGQLDRASSSLYVAFSTCVIKVPLGRG--ERHG 519
DB 446 DSVVTEELQVDF-----PGVPVENLYVVRMCGDSKLVVVSDEILAKLHRGCKKIT 499
QY 520 KCKKTCIASRDPYCGWKKEGAGCSHL-SPN---SRLTFEGDIERGNTDGLGDCHNSFVAL 575
DB 500 NCRF-CVSLQDFPYCANDNVELKCTAVGSPDKWACKGRFTQNISJGEHKACGRFPCTIV- 557
QY 576 NCHSSLLPSTTTSDSTA-----QEGYESRGMGLDKWKLHLSPTDPLGAVSSHNHQ 628
DB 558 ---ASPVTGPTTKSSGCPVHSIQAEFEPE---IDNE-VIGVDDSWIPNTLAEINHA 610
QY 629 DKKGVIRESYLKGHDQLVVPV---TL-LAI-----AVILAFYVGVAFS 666
DB 611 GSK-----LPSSQEKLPYTAETLTIAIVTSLGALVGVF-SGFLFS 652
```

## RESULT 8

US-08-833-391-58

Sequence 58, Application US/58033391

Patent No. 631578;

## GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

APPLICANT: Klocklin, Alex J.

APPLICANT: Matthews, David

APPLICANT: Bentley, David R.

APPLICANT: O'Connor, Timothy

TITLE OF INVENTION: the Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESS: SCIENCE &amp; TECHNOLOGY LAW GROUP

STREET: 268 Bush Street, Suite 1200

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MED'UM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent'n Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/833,391

```
? FILING DATE:
? CLASSIFICATION: 53C
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/121,713
? FILING DATE: 13-SEP-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Osman, Richard A.
? REGISTRATION NUMBER: 36,629
? REFERENCE/DOCKET NUMBER: H94-C02-1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415)343-4341
? TELEFAX: (415) 343-4342
? TELEX:
? INFORMATION FOR SEQ ID NO: 58:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 730 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08 833-391-58
```

```
Query Match 17.8%, Score 969.5; DB 3; Length 730;
Best Local Similarity 34.4%, Pred. No. 4,5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;
```

```
QY 8 LYFTLLHFACGPEDESEPISSHGNYTKQYFVGHKPGRNITQR-----HRLDIQMI 61
DB 11 LKVALH--AAANQNVSP-----KMYVQF-----QSERVQRPFGNESHKQHFLL 54
QY 62 MINNGTLVIAARSHIYVTD--TSHTTEIYCSKKLTWKERQADVTCRMKGKHKDECHN 119
DB 55 EKDNHSLLVGARNVYVNISRLUTEFEQ-----RIEWHSSGARRELOVYKKGSEDDCON 109
QY 120 FIKVLLKNDALFVCGTNAFNSCRNYKMDTEPFC-----EFSQWACRCPYDAKHANV 174
DB 110 YIRVLAKIDDERVLCGTNAVYKPLCRHVALKD---GDVVVEKEVEGRLGCLPDPDCHNST 165
QY 175 ALPADGKLYSATVDFLAIDAVIYRSLGESPTLTUVKHSKWLKEPYVQAVDGYIYF 234
DB 166 AYYSEGQYSATVADFSGTDFLIYRG-----PLRTSRDLKQLNAPNFVNTMEYNDPIFF 220
QY 235 FFRFIAVEYNTMGVPRVAQVCKNDMGSGQRVLEKQWTSFLKARLNCVPGDSHYFN 294
DB 221 FFRFIAVEYNTMGKAIYSRVARVCKDKGPHQGGDR-WTSFLKSRJLNCVPGDYPFFN 279
QY 295 ILQAVTVDIRIN-GRDY---VLATFTSPYNSIPGSAVCAVDYDVLASVFTGRFKEQKSPD 350
DB 290 BIQSTSDITBGNVGGQVEKLIYGVFTTPVNSIGSSAVCAFSMKSLIESFDGPFKEQETMN 339
QY 341 STWTFVPRDERVPRKRCGCCAGSSSLERYATSNBEPDCLNFIKTHPLMDRAVPSIRNPW 410
DB 340 SKWLAVPSCKVPEPRGQCVND-----SRTLPDVSVNPFVKSHTLMDRAVPAFFTRPI 391
QY 411 FLRTWVRVRLTKIAVD---TAAGFYQNHVTVFVGSSEKGIILKFLARIGNSGFJN----- 461
DB 332 LIR-SLQVRFTKAVQCVQVATPDG--KAYDVLFIQTGGKVIKAL-----NSASFDSDTV 445
QY 462 DSJFLEMSVYNSEKSYDGVDERKRMNGQLDRASSSLYVAFSTCVIKVPLGRG--ERHG 519
DB 446 DSVVTEELQVDF-----PGVPVENLYVVRMCGDSKLVVVSDEILAKLHRGCKKIT 499
QY 520 KCKKTCIASRDPYCGWKKEGAGCSHL-SPN---SRLTFEGDIERGNTDGLGDCHNSFVAL 575
DB 500 NCRF-CVSLQDFPYCANDNVELKCTAVGSPDKWACKGRFTQNISJGEHKACGRFPCTIV- 557
QY 576 NCHSSLLPSTTTSDSTA-----QEGYESRGMGLDKWKLHLSPTDPLGAVSSHNHQ 628
DB 558 ---ASPVTGPTTKSSGCPVHSIQAEFEPE---IDNE-VIGVDDSWIPNTLAEINHA 610
QY 629 DKKGVIRESYLKGHDQLVVPV---TL-LAI-----AVILAFYVGVAFS 666
DB 611 GSK-----LPSSQEKLPYTAETLTIAIVTSLGALVGVF-SGFLFS 652
```

```

RESULT 9
US-09-856-610-58
: Sequence 58, Application US/090606:0
: Patent No. 6344544
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.
: APPLICANT: Matthes, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: The Semaphorin Gene Family
: NUMBER OF SEQUENCES: 120
: CORRESPONDENCE ADDRESS:
: ADDRESSES: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 Bush Street, Suite 3200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-PCS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/060,610
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/035,269
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: B94-002-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 343-4341
: TELEFAX: (415) 343-4342
: TELEX:
: INFORMATION FOR SEQ ID NO: 58:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 730 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-856-610-58

Query Match 17.8%; Score 969.5; DB 4; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LYFTLLHFAGAGFPESEP-SISHGNYTKQYVVFVGHKPGRNTTOR-----HRLDIQMI 61
DB 11 LLWVALH--AAAVNDVSP-----KNVQCF-----GEERVQRFGLKSHKQHEKLL 54
QY 62 MINMGILYIARCHIYIVDID--TSHTEEVYCSKLTWRSQRADVDTCKMKGKDECHN 119
DB 55 EKDHNSLLVGARIVNISIRJCUTEFTEQ-----RIEWHSSGAERELCYLKGSEDDQK 109
QY 120 FIKVLKKNDDALFVCGTNAFNSCRNKKYKDTLEPFGD-----EFGMARCPYDAKHANV 174
DB 110 YIRVLAKIDDRVLICGTNAYKPLCHRYALKD---GCVVWEYEGRGGLCPDPDHNST 165
QY 175 ALFADGKLYSATVTDPLAIDAVIYRS-GHSPTLRTVYHDSKWKKEBYFVOAVDYGDIYVF 234
DB 166 AIVSEGLYSATVADESSTDELIYRG-----PLRTERSDLKQLNAPNFVNTNEYNDIFF 220
QY 235 FFEIEINVEYNTGKVIYFPRVACVNDMGSCORVLEKNTFTFKALNCSVPGDSHFYFN 294
DB 221 FFEIETAYEINGKAIYRSVARVKDKGPHGGGR-WTSFLKSLNENVEYQYFFYN 279
QY 295 ILOAVTDVIRIK-GRDV---VLATFTSPYKSIQSSAVCAVDNLDIASVFTGFEKCKSPC 350

```

```
50 11 LKVALH--BAAWNVDP-----KNYVDF-----GEEVQVFLGNESKZHFLL 54
QY 62 YIMNGTLYIAARHIYVDID--TSHTERYGSKKLTWMSRQAVDTQRYKGRHKBCHN 119
Db 55 EKOHNSLUGARNIVYINISLRLTEPTEQ-----R-EMSSGAPRELCVYKKGSEDCCN 109
QY 120 FIKVLKKRDALFVCGTNAPNSCANYKDTSEFEG-----EFSCMARCPCYAKHAY 174
Db 110 YIRVLAKIDSRVLICGTNAYKLCRYALK-----GVYVXEYEGRLCPDFDHST 169
QY 175 ALFADKLYSATVDFIALDAVIYRSLGESPTQYVKEHSEKMKKEPYPCAVDYGVIYF 214
Db 166 AIVSEGLYATVACFSGTDL-YRG-----P-RTESLKLCLNAPVYVWYNEVDPIF 229
QY 235 FFEIAYEYTNKVVPPRPAQVCKNDMGGSQVLEKOWTSPFKALNCSVPCDSFEYFN 294
Db 221 FFEIAYEYTNKKAIVSRVAPVCKEKGPHOQGH-WTSPFKSPKNSVNDYDFEYFN 279
QY 295 ILOAVTGVIRN-GRV-----VLATESPNSIRGSAVCAVDMFIASVUTGKRECKSPD 380
Db 280 EIGSTSLICGKGGGVKLYGVFTVFNISIGSAVCAFPKYSKILSEDPGGFKGCEYTM 379
QY 351 STWTFDEVERPRFOCCAGSSSJERYATSNREFFDCTIANFIMTHPLMDPAVPSIFNRP 410
Db 340 SNWLAVPSLKVPRPPQCVND-----SR-LPQVSNVYKWSHTLYDSAVPAFTPI 391
QY 411 FLRTWYRLTKIAYD-----TAAGPQNVHTVFLGSEKGLILKFIARIGNSFLN----- 461
Db 392 LIRISLQYRFTKIAVQOQVTPQ--KAVDV-PIGTCDKVKIKAL-----NSASFSSQTV 443
QY 462 DSLFLEBMSVYNSEKSYDGVESKEMGMQLDRASSLYVAESTCVIKVPLGR--ERHG 519
Db 446 DSVVIEBLQV-P-----PGVPKNLYVYRNDGDSKLWVYSDDE-LAIKLRCSCKT 490
QY 520 KKKTKTIASRDPYGMWKEGACSHL-SPN---SR-LTFQCI-ERKNTDGLGCHNSFVAL 575
Db 500 NCRB-CVSLQCPYCAMDNVELKCTAVGSPWSAGKRRFIQNTSLGKHKACGGRPOTEIV 557
QY 576 NGRSSSLPSTTTSDSA-----QGVESRGNLDMKHLSDPSTDFLGAVSSHHC 628
Db 558 -----ASVPFQPIKSGGDPVHS:HQAEFEPE---IDNEIVIGDSDSKV:PNTLABINHA 610
QY 629 DKGVIRESYKXGHDQGVV-----L-LAI-----AVLAFYNGAVFS 666
Db 611 GSK-----JFSQSKLPYITAEITIA:VTSCLGALVGFISGFIFS 692
```

```
RESULT 11
US-08-121-713D-60
; Sequence 60, Application US/56121730
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
```

```
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osmar, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4342
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-121-713D-60

Query Match 15.18; Score 880; DB 1; Length 650;
Best Local Similarity 35.84; Proc. No. 1.3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

QY 114 KQCHNFIKVLKKNDALFVCGTNAPNSCANYKMD-----TLEPPCEBFSGMARCPYDA 169
Db 1 EDCQNVIRIWPSPGSLFVCGTNSFRPNCNTYIIISDSNYTLEA---TKXGQAVCPYDP 57
QY 170 KHANVALFADOKLYSATVDFLDAIDAVIYRSLGESPTQYVKEHSEKMKKEPYPCAVDYG 229
Db 58 RHNSTSVLAONELYSGTAVDFSGSDPIIYRE-----PLQTEQYDLSLSLNAPNFVSSTQG 112
QY 230 DYIYFFPREZAVENYTKGVYFPRVACVCKNDMGGSQVLEKOWTSPFKALNCSVPGDS 289
Db 113 DFVYFFPREZAVENYTKGVYFPRVACVCKNDMGGSQVLEKOWTSPFKALNCSVPGDS 289
QY 290 HFYFNILQAVTDVIR-----INGREVLIAETSTPNSIPGSAVCAVDMLOIASVFTGRKE 345
Db 172 PFYFNEIOSASNLVEGOVGSMSKLIYGVFNTPSNIPGSAVCAVCAFALOD-ADTFEGQFKE 231
QY 346 QKSPDSQVTPVDERVPRKPRGCCAGSSJERYATSNREFFDCTIANFIMTHPLMDPAVPSI 425
Db 232 QTG-NSNMLPYNNAKVDPDRPGS-----HNSRALPDTLNFIKTSLMDENYPAF 283
QY 406 FNRWFLRTVYRVELTKIAYD---TAAGPYQVHTVYFLGSEKGLILKFIARIGNSGFL 460
Db 284 PSQILVIRTSTIYFTQIAYDAQIKTPG--KYCVIFVGTGDKGIKISVNASADSADK 341
QY 461 NCSLFLEBMSVYNSEKSYDGVESKEMGMQLDRASSLYVAESTCVIKVPLGR 519
Db 342 VTSVVEIEDVLTYS-----SPIRNLVIRVTQVCPKQGSYDCKGLIIVTDSQVA-CLH 397
QY 514 RC--VRHCKCKTKTIASRDPYGMWKEGAC-SHL-SPN---SR-LTFQCI-ERKNTDGLGDC 568
Db 396 RCHNDK-TSCSE-CVALCPYCAMDNVELKCTAVGSPWSAGKRRFIQNTSLGKHKACGGRPOTEIV 449
QY 569 HNSFVALNGHSSSLPSTTTSDSAQGVESRGNLDMKHLSDPSTDFLGAVSSHHC 628
Db 450 -----HACPSGKINSKCANAGECKGFRNDM-----DLIDS-----RRQ 483
QY 629 DKGVIRESYKXGHDQGVV-----L-LAI-----AVLAFYNGAVFS 666
Db 454 SKQGEIIDKIDKNFEDIINAQYTVETLWMAVLGASIFS 521
```

```
RESULT 12
US-08-935-258-60
; Sequence 60, Application US/58835268
; Patent No. 5927826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
```

;; TITLE OF INVENTION: The Semaphorin Gene Family

;; NUMBER OF SEQUENCES: 100

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

;; STREET: 268 Bush Street, Suite 3200

;; CITY: San Francisco

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 94104

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/619,266

;; FILING DATE:

;; CLASSIFICATION:

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US/08/121,713

;; FILING DATE: 13-SEP-1993

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Osman, Richard A.

;; REGISTRATION NUMBER: 36,607

;; REFERENCE/DOCKET NUMBER: B94-002-1

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (415)343-4342

;; TELEFAX: (415) 343-4342

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 60:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 650 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-08-835-268-60

Query Match 16.1%, Score 880; DB 1; Length 650;

Best Local Similarity 35.8%; Pred. No. 1.3e-74;

Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

QY 114 KDCRPFKVKLLKKNDALFVCGTNAFNPSCRNYKMD---TLEPFGDSEFGMARCPYDA 169

DB 1 EDCQYIRIMVFPGRLEFVCGTNSFRPMCNYYIISDSNYTLEA---TKNGQAVCPYDP 57

QY 170 KHANVALPAGKLYSATVDTFLAIDAVYKSLGESPTLRVTKHDSKWLKBPYFQAVDYG 229

DB 58 RHNSTSVLADNELYSGTVADESGSDPIIYRE-----PLQTEQYDSLSLNAPNFVSSFTQG 112

QY 230 DYIYFFPRETAVENTMGKVYFPRVAQVCKNDXGSGORVLEKQWTSFLKARLNCSPGDS 289

DB 113 DFVYFFPRETAVFVFCNGKAIYRVARVCKWDKGGPHR-FRNRKTSFLKSLKNCSPGQY 171

QY 290 HFYFNILQAVTDVIR-----INGRDVWLATFTPTPNSIPGSAVCAYDMLDIASVFTGRFKE 345

DB 172 PFYFNE-QSASNLVEGOYQSGMSKLVGVPTFSNIPGSAVCAPALQDIADTFEQPVE 231

QY 346 QKSFDSWTVPYDERVFKPFGCAGSSSLERATSNRPDCLINIKTHPLVDENAFSI 405

DB 232 QTSINSXWLPVNAKVPDPFRGSC-----HNDGRALPDEFINFKTHSLMCDENYFAP 283

QY 406 FNRWEVTRVYRLTRIADV---TAAGPYQKHVVVLGSEKGIILKFL-ARIGNSGFL 460

DB 284 FSQILVTRSTVIRFTQIANDQIKITGG--KTYDVIFVUTDGHKIKSNVAESASASAK 341

QY 461 NCSLFEEMSYNKEKSGYDGVGDKRM-QNQLDRASSSLY-----VAFSTQVIKVPFG 513

DB 342 VTSVVVEIDVLTNS-----EPINKLEVRVYQYDQKSGYDQKGLIIVTDSQVAIQH 397

QY 514 RC--ERAGKCKKTCAGSDPYCQWKEGGAC-SHLSFN--SPITFQCIERKNTGSLGDC 569

DB 398 RCHNKDKITGSC-QVA-QEYCAWDKIAGKCRSEGAERWTEENTFYCNVATGQ-----449

QY 569 HNSFVALNCHSSSLSTTTSDSTAEGYESRGMGLDWKHLIDSPQSTDELGAIVSSHNO 628

DB 450 -----HAACFSKINSKDMAGEQKGRNDM-----CLLS-----RRQ 4E3

QY 629 DKKGVIRESYLKGDQLVPVTLTAIAVILAFVYGVAFS 666

DB 484 SKCEILDNICKNFEDLNKQYVETILWAVLAGSIFS 521

RESULT 13

US-09-060-692-60

;; Sequence 60: Application US/09060692

;; Patent No. 5335965

;; GENERAL INFORMATION:

;; APPLICANT: Goodman, Corey S.

;; APPLICANT: Kolodkin, Alex L.

;; APPLICANT: Matthes, David

;; APPLICANT: Bentley, David R.

;; APPLICANT: O'Connor, Timothy

;; TITLE OF INVENTION: The Semaphorin Gene Family

;; NUMBER OF SEQUENCES: 100

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

;; STREET: 268 Bush Street, Suite 3200

;; CITY: San Francisco

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 94104

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/060,692

;; FILING DATE:

;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US/08/121,713

;; FILING DATE: 13-SEP-1993

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Osman, Richard A.

;; REGISTRATION NUMBER: 36,607

;; REFERENCE/DOCKET NUMBER: B94-002-1

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (415)343-4341

;; TELEFAX: (415) 343-4342

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 60:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 650 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-09-060-692-60

Query Match 16.1%, Score 880; DB 2; Length 650;

Best Local Similarity 35.8%; Pred. No. 1.3e-74;

Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

QY 114 KDCRPFKVKLLKKNDALFVCGTNAFNPSCRNYKMD---TLEPFGDSEFGMARCPYDA 169

DB 1 EDCQYIRIMVFPGRLEFVCGTNSFRPMCNYYIISDSNYTLEA---TKNGQAVCPYDP 57

QY 170 KHANVALPAGKLYSATVDTFLAIDAVYKSLGESPTLRVTKHDSKWLKBPYFQAVDYG 229

DB 58 RHNSTSVLADNELYSGTVADESGSDPIIYRE-----PLQTEQYDSLSLNAPNFVSSFTQG 112

QY 230 DYIYFFPRETAVENTMGKVYFPRVAQVCKNDXGSGORVLEKQWTSFLKARLNCSPGDS 289

DB 113 DFVYFFPRETAVFVFCNGKAIYRVARVCKWDKGGPHR-FRNRKTSFLKSLKNCSPGQY 171

QY 290 HFYFNILQAVTDVIR-----INGRDVWLATFTPTPNSIPGSAVCAYDMLDIASVFTGRFKE 345





```

/
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 58/835,268
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Osman, Richard A.
/ REGISTRATION NUMBER: 36,627
/ REFERENCE/DOCKET NUMBER: B94-002-1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415)343-4341
/ TELEFAX: (415) 343-4342
/ TELEX:
/
/ INFORMATION FOR SEQ ID NO: 60:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 650 amino acids
/ TYPE: amine acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-09-060-610-60

Query Match 16.1%; Score 860; DB 4; Length 650;
Best Local Similarity 35.6%; Pred. No. 1,3e 74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 13;

QY 1-4 KKECHNF:KVLKKNKDCALFVGSTNAENPSCNRYQND-----TLRPFQDEFSQARCPYDA 169
DDB 1 EDCQNYRIMVVPFGRLFVGTNTSFRPMCTYIISDSNYTLEA---TKGCAVCPYDP 57
QY 170 KIANVALFADGKLYANVTTFIAIDAVIYRSGESPTLRTYKHSKWKKEPYFOAVDYG 229
DDB 56 RHNS:SVLADNELYSGTVADPSGDFIARE-----PLQEQYDLSLNAPNFUSSTQD 112
QY 230 DYIYFFPFBIAVEYNTMGKVPFPAQVCKNDGSGORVLEKQNTSPFKARINSVFQDS 289
DDB 113 DEVFYFPECTAVEFTGCKAIVRYRVARVCKMDKGGPHR-FENRMTSPFKASLNCSTPGDY 17;
QY 290 HYFNFILCAVTDVIR-----LNGKDVLTATESPVNYSIFGGAVCAQMDLAIASVFTGRKE 345
DDB 172 PFFNFIEQASNLVGGVGSMSKLYGVNTPSHSIFGSAVCAALACIADTFEGQKE 231
QY 346 QKSPSTWTPVLEVPKPPRPGCCAGSSSLERYATNRPDITLNFKTHPIMTEAVTSI 405
DDB 232 CTGKNSMILPNNAKVPOPRPGSC-----HNSRALPPTLNPKTHSLYDENVAF 283
QY 406 ENRPWELRTMAYRLTKIAC-----TAAGPQNHTVYFLGSEKGIILKFL-ARIONSGEL 460
DDB 284 FSPQ-LVTRISTIVRFITQIADACIKTPGG--KTYDVFVGTGTHGKIISKVAESACSAK 341
QY 461 NDSJFLEMSVYNSEKSYDGVDEKRLX-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
DDB 342 VTSVVEIEDVLTKS-----EPTRNLEIVRTNOYDQPKDGSYDEGKLIIVTDSQVVALCLH 397
QY 514 RC--ERHCKKXTCIASHDPYCGMIKEGGAC-SHLSPN--SLTFEQDIERGNTDGLGDC 569
DDB 398 RCHNDKITSCE-SVALQDPYCAWQKIAGKCRSHGAPRWLEENFYQKAVATGQ----- 449
QY 569 HNSFVALNCHSSSLIPSTTSDSTAGEVSRGMDKWKHLLDSPDSTDPLGAVSSHQ 626
DDB 450 -----HAACPSGKINSKANAGECKGRNDN-----DLLES-----PRQ 483
QY 629 DKKGVIRESYLKGHCLVPVTLIAIVLAFVMGAVES 666
DDB 484 SKDQEIIONIDKNFEDIINAQYTVETLVAVLAGSIPS 521
```

Search completed: October 23, 2003, 17:10:59

Job time : 23 secs



```

PF 26-NOV-1999; 99WO-EP09215.
XX
XX
XX 26-NOV-1998; 98EP-0122411.
XX
XX (PLAC ) MAX PLANCK GES FÖRDERUNG WISSENSCHAFTEN.
XX
XX Behl C; Klostermann A;
XX
XX WPI: 2000-40065/34.
XX
XX P-PSDB; MAY7461.
XX
XX Nucleic acid coding for human semaphorin 6A-1, used as diagnostic agent.
XX
XX therapeutic agent, for modulating immune system, in gene therapy or for
XX
XX effecting differentiation, cytoskeletal stabilization and/or plasticity
XX
XX
XX Claim 2; Page 21; 53pp; English.
XX
XX The present sequence is a DNA encoding binding domain of transmembraneous
XX
XX human semaphorin 6A-1 ((HSA)SEMA6A-1) which is involved in neuronal
XX
XX development and regeneration mechanisms during apoptosis. The binding
XX
XX domain shows homology to zyxin protein and selectively binds to members
XX
XX of Ena/VASP protein family, especially Evi. (HSA)SEMA6A-1 is a
XX
XX member of protein family displaying secreted or transmembrane-based
XX
XX repulsive guidance cues critically involved in neuronal development.
XX
XX Expression of (HSA)SEMA6A-1 is highest in embryonic brain and
XX
XX kidney and moderate in lung. The present sequence is useful as diagnostic
XX
XX and therapeutic agents, for modulating the immune system, in gene
XX
XX therapy, for effecting differentiation, cytoskeletal stabilisation
XX
XX and plasticity.
XX
XX Sequence 216 BP; 45 A; 85 C; 51 G; 35 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 216; DB 21; Length 216;
XX
XX Best Local Similarity 100.0%; Pred. No. 2e-52;
XX
XX Matches 216; Conservative 0; Mismatches 0; Indels 0; Caps 0;
XX
XX
XX QY 1 CCGCGCGCGCGCGCGAGGGTGGAGTCCAGTCCAGTGCACAGTCCAGGCATCTGGC 60
XX
XX DB 1 CCGCGCGCGCGCGCGAGGGTGGAGTCCAGTCCAGTGCACAGTCCAGGCATCTGGC 60
XX
XX
XX QY 61 CAGCGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAAGCGCTACAGTCACTGCAAGGTGC 120
XX
XX DB 61 CAGCGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAAGCGCTACAGTCACTGCAAGGTGC 120
XX
XX
XX QY 121 GGGGTGAAGCGTACGGCTGTGGTAAAGCGGAGTACGCGCCAAAGCTCCCTTTGGTGGC 180
XX
XX DB 121 GGGGTGAAGCGTACGGCTGTGGTAAAGCGGAGTACGCGCCAAAGCTCCCTTTGGTGGC 180
XX
XX
XX QY 181 CTTCACATCCATGAAGCGGCATATGATGGGTGACA 216
XX
XX DB 181 CTTCACATCCATGAAGCGGCATATGATGGGTGACA 216
XX
XX
XX RESULT 2
XX
XX AAC98050
XX
XX ID AAC98050 standard; cDNA; 1472 BP.
XX
XX AC AAC98050;
XX
XX
XX D7 09-MAR-2001 (first entry)
XX
XX
XX DE Human colon cancer antigen nucleotide sequence SEQ ID NO:6.
XX
XX
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX
XX identification; cytoskeletal; cardiovascular; neuroprotective; vulnary;
XX
XX immunomodulatory; muscular; gynaecological; gastrointestinal;
XX
XX nephrotropic; antiinfective; antibacterial; gene therapy; wound;
XX
XX neural disorder; immune system disorder; muscular disorder;
XX
XX reproductive disorder; gastrointestinal disorder; renal disorder;
XX
XX infectious disease; cardiovascular disorder; ss.
XX
XX
XX OS Homo sapiens.

```

```

XX
XX WC2000055551-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05683.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI: 2000-587534/55.
XX
XX P-PSDB; AAR53293.
XX
XX Colon cancer associated gene sequences, referred to as colon cancer
XX
XX antigens, useful for the treatment, prevention, and diagnosis of colon
XX
XX disorders such as colon cancer.
XX
XX Claim 1; Page 510-511; 2104pp; English.
XX
XX AAC9799; 10 AAC98763 encode the human colon cancer associated proteins,
XX
XX called human colon cancer antigens, given in AAB53234 to AAB54006. The
XX
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;
XX
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
XX
XX vulnary, nephrotropic, antiinfective and antibacterial activities, and
XX
XX can be used in gene therapy. The colon cancer antigen polynucleotides,
XX
XX proteins and antibodies to the proteins are useful for the prevention,
XX
XX treatment and diagnosis of colon disorders, such as colon cancer. The
XX
XX polynucleotides may be used in diagnostics and research, such as for
XX
XX chromosome identification, and as hybridisation probes. The proteins
XX
XX may also be used to prevent diseases such as neural disorders, immune
XX
XX system disorders, muscular disorders, reproductive disorders,
XX
XX gastrointestinal disorders, wounds, renal disorders, infectious
XX
XX diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
XX
XX AAB54007 represent sequences used in the exemplification of the present
XX
XX invention.
XX
XX Sequence 1472 BP; 447 A; 431 C; 299 G; 301 T; 4 other;
XX
XX
XX Query Match 100.0%; Score 216; DB 21; Length 1472;
XX
XX Best Local Similarity 100.0%; Pred. No. 3.2e-52;
XX
XX Matches 216; Conservative 0; Mismatches 0; Indels 0; Caps 0;
XX
XX
XX QY 1 CCGCGCGCGCGCGCGAGGGTGGAGTCCAGTCCAGTGCACAGTCCAGGCATCTGGC 60
XX
XX DB 1 CCGCGCGCGCGCGCGAGGGTGGAGTCCAGTCCAGTGCACAGTCCAGGCATCTGGC 60
XX
XX
XX QY 61 CAGCGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAAGCGCTACAGTCACTGCAAGGTGC 120
XX
XX DB 61 CAGCGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAAGCGCTACAGTCACTGCAAGGTGC 120
XX
XX
XX QY 121 GGGGTGAAGCGTACGGCTGTGGTAAAGCGGAGTACGCGCCAAAGCTCCCTTTGGTGGC 180
XX
XX DB 121 GGGGTGAAGCGTACGGCTGTGGTAAAGCGGAGTACGCGCCAAAGCTCCCTTTGGTGGC 180
XX
XX
XX QY 181 CTTCACATCCATGAAGCGGCATATGATGGGTGACA 216
XX
XX DB 181 CTTCACATCCATGAAGCGGCATATGATGGGTGACA 216
XX
XX
XX RESULT 3
XX
XX AAR13995
XX
XX AAR13995 standard; cDNA; 2123 BP.
XX
XX AC AAR13995;
XX
XX
XX D7 26-JUN-2001 (first entry)
XX
XX
XX DE Human cDNA sequence SEQ ID NO:11072.
XX
XX
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

```







DR WPI: 2001-639362/73.  
DR P-PSDB: ABG04066.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1: SEQ ID No 4057; 103pp; English.  
PS  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AAS64197-AAS94564 represent novel human  
XX diagnostic coding sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 2592 BP; 655 A; 754 C; 659 G; 524 T; 0 other;  
Query Match 100.0%; Score 216; DB 23; Length 2592;  
Best Local Similarity 100.0%; Pred. No. 3.7e-52;  
Matches 216; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGCGCGCCGCGCCGCGAGAGGGTGACCTCCATCCAGGTGCACAGCTCCAGCCCATCTGGC 60  
DB 2374 CGCGCGCCGCGCCGCGAGAGGGTGACCTCCATCCAGGTGCACAGCTCCAGCCCATCTGGC 2433  
QY 61 CAGCGCGTGACGTCTGAGCGAGCCAGCCGCTCAAGCCCTACACACTGACAGAGTGG 125  
DB 2434 CAGCGCGTGACGTCTGAGCGAGCCAGCCGCTCAAGCCCTACACACTGACAGAGTGG 2493  
QY 121 GGGCTGAAGCGTACGCCCTCGCTAAAGTCGGAGGTACGCCCAAGCCATCCTTGGTCCG 180  
DB 2494 GGGCTGAAGCGTACGCCCTCGCTAAAGTCGGAGGTACGCCCAAGCCATCCTTGGTCCG 2553  
QY 181 GTTCCACATCCATGAAGCGCAATGATGGTGATACA 216  
DB 2554 GTTCCACATCCATGAAGCGCAATGATGGTGATACA 2580  
RESULT 9  
ID AAS68807 standard3: cDNA: 3039 BP.  
XX  
AC AAS68807;  
XX  
XX 13-FEB-2002 (first entry)  
DE DNA encoding novel human diagnostic protein #4611.  
XX  
KW human; chromosome mapping; gene mapping; gene therapy; forensics;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200275057-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US86631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (BYSE-) HVSFQ INC.  
XX  
XX Drenaco RT, Liu C, Tang Y;  
XX  
XX WPI: 2001-639362/73.  
XX P-PSDB: ABG04620.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 1: SEQ ID No 4611; 103pp; English.  
PS  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AAS64197-AAS94564 represent novel human  
XX diagnostic coding sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 3039 BP; 741 A; 869 C; 781 G; 648 T; 0 other;  
Query Match 100.0%; Score 216; DB 23; Length 3039;  
Best Local Similarity 100.0%; Pred. No. 3.8e-52;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGCGCGCCGCGCCGCGAGAGGGTGACCTCCATCCAGGTGCACAGCTCCAGCCCATCTGGC 60  
DB 2821 CGCGCGCCGCGCCGCGAGAGGGTGACCTCCATCCAGGTGCACAGCTCCAGCCCATCTGGC 2880  
QY 61 CAGCGCGTGACGTCTGAGCGAGCCAGCCGCTCAAGCCCTACACACTGACAGAGTGG 120  
DB 2881 CAGCGCGTGACGTCTGAGCGAGCCAGCCGCTCAAGCCCTACACACTGACAGAGTGG 2940  
QY 121 GGGCTGAAGCGTACGCCCTCGCTAAAGTCGGAGGTACGCCCAAGCCATCCTTGGTCC 180  
DB 2941 GGGCTGAAGCGTACGCCCTCGCTAAAGTCGGAGGTACGCCCAAGCCATCCTTGGTCC 3000  
QY 181 GTTCCACATCCATGAAGCGCAATGATGGTGATACA 216  
DB 3001 GTTCCACATCCATGAAGCGCAATGATGGTGATACA 3035  
RESULT 10  
ID AAS69721  
XX  
XX AAS69721 standard: cDNA: 3039 BP.  
XX  
AC AAS69721;  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX



DE DNA encoding novel human diagnostic protein #25525.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2002; 2001US-0540217.

XX 23-AUG-2000; 2000US-0549167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YJ;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG25534.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX Claim 1: SEQ ID NO 25525; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (II) is useful in gene therapy techniques  
CC to restore normal activity of (I) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical,  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA564197-AA594564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pat\_sequences.

XX Sequence 3039 BP; 741 A; 865 C; 781 G; 642 T; 0 other;

XX Query Match 100.0%; Score 216; DB 23; Length 3039;  
XX Best Local Similarity 100.0%; Pred. No. 3.8e-52;  
XX Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCGCGCGCGAGAGGGTGGACTCCAGCTGCACAGCTCCAGCCACCTGGC 60  
DB 2821 CGCGCGCGCGCGCGAGAGGGTGGACTCCAGCTGCACAGCTCCAGCCACCTGGC 2800

QY 61 CAGGCGGTGACTGTCTCGAGCGAGCGCCAGCGCTCAACGCTCAACTCAGTCAGCAAGTGG 120  
DB 2881 CAGGCGGTGACTGTCTCGAGCGAGCGCCAGCGCTCAACGCTCAACTCAGTCAGCAAGTGG 2840

QY 121 GGGGTGAGAGGTACCGCTTCCTTAAGCGGAGCTACCGCCCAACACCATCTTTCCTGCC 180  
DB 2941 GGGGTGAGAGGTACCGCTTCCTTAAGCGGAGCTACCGCCCAACACCATCTTTCCTGCC 3000

QY 18: GTTCCACATCCATGAGCCCAATCATCGCTGTACA 216  
DB 3001 GTTCCACATCCATGAGCCCAATCATCGCTGTACA 3036

RESULT 11  
AAH17625  
ID AAH17625 standard; cDNA; 3041 BP.

XX AC AAH17625;

XX 24-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:17153.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0100253.

XX 31-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI) HELIX RES INST.

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama I, Wakanatsu A, Nagai K, Otsuki I;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX Claim 8: SEQ ID 17153; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 3041 BP; 804 A; 886 C; 714 G; 637 T; 0 other;

XX Query Match 100.0%; Score 216; DB 22; Length 3041;  
XX Best Local Similarity 100.0%; Pred. No. 3.8e-52;  
XX Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCGCGCGCGAGAGGGTGGACTCCAGCTGCACAGCTCCAGCCACCTGGC 60  
DB 3001 CGCGCGCGCGCGCGAGAGGGTGGACTCCAGCTGCACAGCTCCAGCCACCTGGC 60





[illegible]

Result No.	Score	Query Match	Length	DB	ID	Description
1	216	100.0	495	12	BM855120	K-EST0137
2	216	100.0	496	9	AV704775	AV704775
3	216	100.0	815	13	BM850016	AGENCOURT
4	216	100.0	862	13	BO723018	BO723018
5	216	100.0	862	13	BO723018	AGENCOURT

495 bp mRNA linear EST 06-MAR-2002  
KX1855120.1 Homo sapiens cDNA clone S22SNJ161-150-D07  
5' mRNA sequence:  
RM855120  
BM855120.1 GI:19211519  
EST.  
Homo sapiens (human);  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 495)  
K.M.Y.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
K.M.Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel.: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 150 row: D column: 07  
High quality sequence stop: 495.









Best Local Similarity 99.5%; Fred. No. 1.8e-47;  
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	CCGCGCGCCGCCCGCCGACAGGGTGGACTCCATCCAGGTGCAGCTCCGACCGCATCTGGC	60
Db	60	CCGCGCGCCGCCCGCCGACAGGGTGGACTCCATCCAGGTGCAGCTCCGACCGCATCTGNC	115
QY	61	CAGGCGGTGACTGTCTGAGGAGCGCCAGCGCTCAAGCGCTACACTCAGTGAAGATCG	120
Db	120	CAGGCGGTGACTGTCTGAGGAGCGCCAGCGCTCAAGCGCTACACTCAGTGAAGATCG	175
QY	121	GGGCTGAAGGCTAGCGGCTCGCTAAAGCGGAGCTACCGCGCAACCATCTTCTGCGC	180
Db	180	GGGCTGAAGGCTAGCGGCTCGCTAAAGCGGAGCTACCGCGCAACCATCTTCTGCGC	235
QY	181	CTTTCACATCCATGAAGCCCAATGATGCGGTGACA	215
Db	240	CTTTCACATCCATGAAGCCCAATGATGCGGTGACA	275

BC032619 6875 bp mRNA linear HTC 04-MAR-2003  
Homo sapiens, similar to sema domain, transmembrane domain (TM),  
and cytoplasmic domain, (semaphorin) 5A, clone IMAGE:5575066, mRNA.  
BC032619  
RC032619.1 GI:22749800  
HTC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
Strausberg, R.  
Direct Submission  
Submitted (06-JUN-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [mgc-help@nih.gov](mailto:mgc-help@nih.gov)  
Tissue Procurement: ATCC/DCT/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc-mgc@nih.gov](mailto:nisc-mgc@nih.gov)  
Akitter, N., Ayala, K., Beckstrom-Sternberg, S., Benjamin, R.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brickley, C., Brooks, S.,  
Chetrich, N., Granits, S., Guan, X., Gupta, J., Hashiguchi, P.,  
Hansen, N., He, S.-L., Karlins, E., Keong, P., Lario, P., Ledas, R.,  
Madore, Q.H., Mastello, C., Maskeri, B., Mastrian, S., McCloskey, J.,  
McDowell, J., Pearson, R., Stantrispop, S., Thomas, P.J., Touchman, S.W.,  
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>  
Series: IRAC plate; 6; Row: 0 Column: 4  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA GI: 11991659  
This clone has the following problem: retained intron.

Location/Qualifiers  
1..6875  
/organism="Homo sapiens"  
/db\_type="mRNA"  
/mol\_type="taxon:9606"  
/clone="IMAGE:5578066"  
/tissue\_type="Skin, melanotic melanoma"  
/clone\_lib="NIH\_MGC\_72"

FEATURES  
source

```

/lab_host="3810B"
/note="Vector: pCMV-SPORT6"
BASE COORDINATE 1784 a 1795 c 1592 g 1704 t
ORIGIN

Query Match 98.5%; Score 212.8; DB 11; Length 6875;
Best Local Similarity 99.1%; Pred. No. 1.7e-46;
Matches 214; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGCCGCCGCCGCCGCCGAGAGGGTGGACTCCATCCAGGTCGACACAGCTCCGAGGCATCTGGC 60
DB 3666 CGCCGCCGCCGCCGCCGAGAGGGTGGACTCCATCCAGGTCGACACAGCTCCGAGGCATCTGGC 3725
QY 61 CAGCGCGTGACTCTCTCGAGGCGACGCCAGCCTCAAGCGCTACAACTCACTGACAAGGTGCG 120
DB 3726 CAGCGCGTGACTCTCTCGAGGCGACGCCAGCCTCAAGCGCTCAAACTCACTGACAAGGTGCG 3785
QY 121 GGCTGAAGGTCAGCGCCCTCGCTAAAGCGGGAGTACCGCCGCCAAACATCCTTTGCTCCG 180
DB 3786 GGCTGTAGGTCAGCGCCCTCGCTAAAGCGGGAGTACCGCCGCCAAACATCCTTTGCTCCG 3845
QY 18: CTTTCACATCCCATGAAGCCCATGATGCGTGATCA 216
DB 3846 CTTTCACATCCCATGAAGCCCAATGATGCGTGATCA 3881

RESULT 9
LOCUS N25829
DEFINITION YAZ2001.r1 Soares melanocyte 2NDHM Homo sapiens cDNA clone
ACCESSION N25829
VERSION N25829.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 462)
AUTHORS Blicher,L., Clark,N., Dubucque,T., Elliston,K., Hawkins,M., Holman,
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M., Parsons,J.,
Rikkila,L., Rohlfing,T., Soares,M., Tan,P., Trevasz,E., Waterston,
R., Williamson,A., Wolfdmatten,P. and Wilson,R.
TITLE The Wasci-Xerox EST Project
JOURNAL Unpublished
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 285 1800
Fax: 314 264 1810
Email: est@watson.wustl.edu
High quality sequence stops: 360
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1117 Std Error: 0.00
Seq primer: 17
High quality sequence stop: 360.
Location/Qualifiers
1..462
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="Gene:3872107"
/db_xref="taxon:9606"
/clone="IMAGE:262465"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="OH10B (ampicillin resistant)"
/clone_lib="Soares melanocyte 2NDHM"
/note="Vector: pT73D (Pharmacia) with a modified
polylinker; Site.1: Not 1; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACATCATCAAGTCGAGCGCGCCAGCTTTTCTTTTCTTTT 3']

```











```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 43280 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURES:
: NAME/KEY: CDS
: LOCATION: 816..14234
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 14351..19945
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 20010..31199
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 31232..36067
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 36249..41774
: US-08-804-227C-1

Query Match 15.3%; Score 35.2; DB 2; Length 43280;
Best Local Similarity 48.5%; Pred. No. 0.78;
Matches 97; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 2 CCGCCGCGCCCGCCGAGGGTGGACTCCATCCAGTGCACAGCTCCAGCAGCTCCAGCATCTGSCC 61
DB 1896 CCGGAGGGGGTGTGCGGGCGTGTGGCGGGGCGGAGCGGGCGGCGAGTGGG 1897
QY 62 AGCGCGTACITCTCGAGGAGCGAGCGAGCGCTCAAGCGCTACACACTCACTGACAGGTGG 121
DB 1836 CGGCTCCAGGGGTGCGGGCGCCCGGTGCGGSHGATGTGAGTCTGAGGTGAGGAGTGG 1897
QY 122 GGTCAAGCTAGCGCCCGTGTAAAGCGGAGGTACCGCGGCAAGCAACCTTGTGTCGG 121
DB 1776 GCGCGGCTGAGCGCGCGCGCGGTAGGGCTGCGGCGAGTACAGCTTCTGCGCTCGC 1713
QY 182 TTTCACATCCATGAAGCC 201
DB 1716 GGTGGGAGTGTGAGGCTC 1697

RESULT 5
US-09-103-840A-2/G
: Sequence 2, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-20007.0;
: CURRENT APPLICATION NUMBER: US/09/103.840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 4403765
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: CDC 1551
: OTHER INFORMATION: "n" bases at various positions throughout the sequence
: OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 15.9%; Score 34.4; DB 3; Length 4403765;
Best Local Similarity 57.4%; Pred. No. 4.2;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

US-09-103-840A-1/G
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103.840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 15.9%; Score 34.4; DB 3; Length 4411529;
Best Local Similarity 57.4%; Pred. No. 4.2;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 22 GIGGACICCATCCAGTGCACAGTCCAGCATCCAGCATCIGGCCAGCGCTGACTGTCTCGAGG 81
DB 744035 GGGATTCCTCTGTAGCGCGCAGCTGTGGCTACCGGGCGCGCGCATCAGCTGGCGG 743976
QY 82 CAGCCCAAGCTCAAGCGGCTACACTCACTACAGTGGGGCTGAAG 129
DB 743975 CGTCTTCACACAGGAGGAGGAGCGCGCGCATCAGCTGGGGTGAAG 743928

RESULT 7
US-09-252-991A-8905/C
: Sequence 5905, Application US/59252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubinfeld et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252.991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 8905
: LENGTH: 846
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8905

Query Match 15.6%; Score 33.8; DB 4; Length 846;
Best Local Similarity 49.2%; Pred. No. 0.72;
Matches 89; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 2 CGCGCGCGCGCGCGCGAGGTGGAGTCCATCCAGGTGCACAGCTCCAGCATCTGSCC 61

```

Db 343 CTTGGCAGCGGACACATCTTTCGGCCCATCGCTTTCGGCGGCGAGGTCGGCGCGG 284  
QY 62 AGCGCGTACTGTCAGGAGGAGCGGAGCGCTACAGGCTACTACAGAGTGG 122  
Db 283 CGTCGGCGCGGAGCGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 224  
QY 122 GGCTGAAGCGTACCGCTCGGTAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 182  
Db 223 CCAGCGAGCGTTTACGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTG 164  
QY 182 T 182  
Db 163 T 163

## RESULT 8

US-09-252-991A-1737  
; Sequence 1737, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1737  
; LENGTH: 1026  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1737

Query Match 15.6%; Score 33.8; DB 4; Length 1026;  
Best Local Similarity 54.4%; Pred. No. 0.76;  
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 62 AGCGCGTACTGTCAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 122  
Db 271 AGCGCGATCAATCAAGGAGTCTTGGCGGAGGCGGAGGAGGAGGAGGAGGAGG 330  
QY 122 GGCTGAAGCGTACCGCTCGGTAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 182  
Db 331 GAAGGAAGAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 390  
QY 182 TTTC 186  
Db 391 TTAC 395

## RESULT 9

US-09-252-991A-9254  
; Sequence 9254, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 9254  
; LENGTH: 2667

; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-9254

Query Match 15.6%; Score 33.8; DB 4; Length 2667;  
Best Local Similarity 49.2%; Pred. No. 0.96;  
Matches 89; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 2 CGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 62  
Db 36 CGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97  
QY 62 AGCGCGTACTGTCAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 121  
Db 96 CGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 157  
QY 122 GGCTGAAGCGTACCGCTCGGTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 181  
Db 158 CCAGCGAGCGTTTACGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTG 217  
QY 182 T 182  
Db 218 T 218

## RESULT 10

US-09-252-991A-1816  
; Sequence 1816, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1816  
; LENGTH: 3228  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1816

Query Match 15.6%; Score 33.8; DB 4; Length 3228;  
Best Local Similarity 54.4%; Pred. No. 1;  
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 62 AGCGCGTACTGTCAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 121  
Db 303 AGCGCGATCAATCAAGGAGTCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 362  
QY 122 GGCTGAAGCGTACCGCTCGGTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 181  
Db 363 GAAGGAAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 422  
QY 182 TTTC 186  
Db 423 TTAC 427

## RESULT 11

US-09-252-991A-2143/c  
; Sequence 2143, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136



```
? CURRENT APPLICATION NUMBER: US/09/252,991A
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/074,768
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 2143
? LENGTH: 3546
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2143

Query Match
Best Local Similarity 54.4%; Score 33.8; DB 4; Length 3546;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 62 AGCGCGTACTCTCTCGATGCGAGCCAGCCCTCAAGCGCTACAACTCACTGAGAGGCTCG 121
DB 3064 AGCGCGGATCAACATCAAGTCTGAGCGAGGCGAGGAGGAGTCTGCGAGTCTGA 3005
QY 122 CGCTCAAGCGTACGCCCTTGGTAAAGCGCGGAGTACGCGCCCAAGCATGCTTTGCTCC 181
DB 3004 GAAGAGAGCGCGGCAAAAGGCGCGCCCTGAGCTGACCTTATCAAGCTGGGAGGCG 2945
QY 182 TTCC 186
DB 2944 TTAC 2943

RESULT 12
US-09-252-991A-7012/c
? Sequence 7012, Application US/09252991A
? Patent No. 6551795
? GENERAL INFORMATION:
? APPLICANT: Marc J. Rubenfield et al.
? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
? TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
? FILE REFERENCE: 107196.176
? CURRENT APPLICATION NUMBER: US/09/252,991A
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/074,768
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 7012
? LENGTH: 630
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7012

Query Match
Best Local Similarity 15.4%; Score 33.2; DB 4; Length 630;
Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 7 CCGCGCCCGCAGAGGTGGAGTCCATCCAGTGCACAGCTCCAGCCATCTGCGAGGCG 66
DB 203 CCGCGCGGATCGAGGTGATCCGCGCATGCGGCGGCGGCGGCGGCGGCGGCGGCG 144
QY 67 GTGACTGTCTCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 126
DB 143 CATCTATGCTGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 64
QY 127 AGCGTACGCCCTC 140
DB 83 AGCGAGTCCGCCCTC 70

RESULT 13
US-09-252-991A-7061
? Sequence 7061, Application US/09252991A
? Patent No. 6551795
? GENERAL INFORMATION:
? APPLICANT: Marc J. Rubenfield et al.
? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
? TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
? FILE REFERENCE: 107196.136
? CURRENT APPLICATION NUMBER: US/09/252,991A
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/074,768
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 7061
? LENGTH: 1548
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7061

Query Match
Best Local Similarity 15.4%; Score 33.2; DB 4; Length 1548;
Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 7 CCGCGCCCGCAGAGGTGGAGTCCATCCAGTGCACAGCTCCAGCCATCTGCGAGGCG 66
DB 623 CCGCGCGGATCGAGGTGATCCGCGCATGCGGCGGCGGCGGCGGCGGCGGCGGCG 682
QY 67 GTGACTGTCTCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 126
DB 683 CATCTATGCTGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 742
QY 127 AGCGTACGCCCTC 140
DB 743 AGCGAGTCCGCCCTC 756
```

## RESULT 15

US-08-993-359-29  
 : Sequence 29, Application US/58693359A  
 : Patent No. 6039942  
 : GENERAL INFORMATION:  
 : APPLICANT: Lassen, Soren P.  
 : APPLICANT: Bech, Lisebeth  
 : APPLICANT: Ohmann, Anders  
 : APPLICANT: Breitholtz, Jess  
 : APPLICANT: Fuglsand, Claus C.  
 : APPLICANT: Ostergaard, Peter F.  
 : TITLE OF INVENTION: Phytase Polypeptides  
 : FILE REFERENCE: 5363-500-15  
 : CURRENT APPLICATION NUMBER: 48/58/993, 359A  
 : CURRENT FILING DATE: 1997-12-18  
 : EARLIER APPLICATION NUMBER: 1480/96  
 : EARLIER FILING DATE: 1996-12-20  
 : EARLIER APPLICATION NUMBER: 1481/96  
 : EARLIER FILING DATE: 1996-12-20  
 : EARLIER APPLICATION NUMBER: 0301/97  
 : EARLIER FILING DATE: 1997-03-18  
 : EARLIER APPLICATION NUMBER: 0523/97  
 : EARLIER FILING DATE: 1997-03-07  
 : EARLIER APPLICATION NUMBER: 1386/97  
 : EARLIER FILING DATE: 1997-12-01  
 : EARLIER APPLICATION NUMBER: 60/346,082  
 : EARLIER FILING DATE: 1997-05-09  
 : NUMBER OF SEQ ID NOS: 32  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 29  
 : LENGTH: 1536  
 : TYPE: DNA  
 : ORGANISM: *Trametes pubescens*  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: (79)...(1407)  
 : NAME/KEY: mat\_peptide  
 : LOCATION: (130)...(1407)  
 : NAME/KEY: sig\_peptide  
 : LOCATION: (79)...(129)  
 US-08-993-359-29

Query Match 15.3%; Score 33; DB 3; Length 1536;  
 Best Local Similarity 53.5%; Pred. No. 1.4;  
 Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY	11	CCCCGCGAGGGTGGAGCTCATCCAGGCGACAGCTCCGAGGCTATCTGGCGGAGGCTGA	70
DB	950	CCGTGCAAGGCGTGGGTATATCAACGAGCTCATCGCGGCTCAGCCGCGCAAGGTGT	1009
QY	71	CTGTCTCGAGGCGCCGAGCTCAACGCTACAACTCAGTCAGCAGGCTGGGGCTGAAGC	130
DB	1010	CCGACACACGCGACGAGGAGCAGCAGCAGCTCGCTCTCCCGGAGACGTTCCGCTCAAC	1369
QY	131	GTACGCCCT	139
DB	1070	GCACGCTCT	1078

Search completed: September 30, 2003, 22:06:17  
 Job time : 27.4914 secs





```

neurod1, pre-cursor cells."
127..1650
/note="unlabeled protein product"
/codon_start=-1
/protein_id="B06551.1"
/db_xref="GI:14942114"
/translation="MSVNSKSGSTGVEHFKRMKMLDRASSLSLVRFSTGVKVPD
GRGRGKKKKICIAHFDHYCWKEGASCHLSPNSRIFEQISRGNTDGGSGCHN
SVFAKGLFHSIRKHGDIQVPTLAIADLAFVGAVFSGTVYGVCDGPRKWAY
VQKAGVLEHSHSRGMSVTKSGDFGDIQSKDKFPAITLPIMHGKLATGNTAKV
LIKADQHLHLLTALPTPESTPTLQGRKPSRGSRWENONIINACTKQHPKMSV
PTDPLKASPSHIPSVVLPITTCQYCHVYVNGKMSVACMALEQAALEVATKE
HLSKSPNHCVALVENLSEIPKVPQKENSILQTFASLSQGLSKPLEMHHSYVDS
YKSYPTNLSLTSKATLTKRNTNSNSHLSKHSQSGRGKQFPAPQVSGVCIQVHS
SQSGGAVTVSCPLSNATNLSITSGILKRITSLKDFEFPKPSAPLISNKPNDQVH
BASE COUNT 550 a 664 c 528 g 381 t
ORIGIN

Query Match 300.0% Score 216; DB 9; Length 2124;
Best Local Similarity 100.0% Pred No. 8,40-42;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCGCGCCCGCCCGCAGAGGGTGAGCTCCATCCAGTTGACACATCCACAGCTCCACAGCATCTCCG
DB 1432 CCGCGCGCCCGCCCGCAGAGGGTGAGCTCCATCCAGTTGACACATCCACAGCTCCACAGCATCTCCG
QY 61 CAGCGCGCTGACTCTCGAGGAGGAGCGACGCTCAACGCGCTACACATCCACAGGACAGCTG
DB 1492 CAGCGCGCTGACTCTCGAGGAGGAGCGACGCTCAACGCGCTACACATCCACAGGAGCTG
QY 121 GGAGTGAAGCGTAGCGGCGGCGCTAAAGCGGAGCTACGCGCTAAAGCGATCTTACTGCG
DB 1552 GAGTGAAGCGTAGCGGCGGCGCTAAAGCGGAGCTACGCGCTAAAGCGATCTTACTGCG
QY 191 CTTTCCATCCGATGAAGCGCGATCAIGAGTGTCAGA 216
DB 1612 CTTTCCATCCGATGAAGCGCGATCAIGAGTGTCAGA 1647

RESULT 4
B0127394
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

OS Homo sapiens (human)
PN JP 2002017375-A/2825
PF 22-JAN-2002
PI 07-JUL-2000 JP 2002253172
PI TOSHIO OIA, TETSUO NISHIKAWA, TAKAO TSOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA.
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/-5, C12N/-19, C12N1/22, C12N5/ PC
10, C12N21/02, C12Q1/58/C12P21/08, G06F17/30, C12N15/00, C12N5/06 CC
Primer for synthesizing full-length cDNA and use thereof. FH Key
Location/Qualifiers
(56). (1741).
FT CDS

```

```

FEATURES
  source
    1..2227
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
BASE COUNT 567 a 736 c 560 g 400 t
ORIGIN
Query Match 100.0% Score 216; DB 5; Length 2227;
Best local Similarity 100.0%; Pred. No. 8.4e-42;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGCGCGCCGCGCGCGAGAGGGTGGACTGCATCCAGCTGCACACACTCCAGGCATCTGCC 60
DB 1226 CCGCGCGCCGCGCGCGAGAGGGTGGACTGCATCCAGCTGCACACACTCCAGGCATCTGCC 1585
QY 61 CAGCGCGTGACTGCTCTGAGGAGCGAGCCAGCGCTCAACGCGCTACAACTCACTGACCAAGTGC 120
DB 1556 CAGCGCGTGACTGCTCTGAGGAGCGAGCCAGCGCTCAACGCGCTACAACTCACTGACCAAGTGC 1645
QY 121 GGGCTGAAGCGTACGGCCCTCGCTAAAGCGGGACGCTACCGCCCAACACCATCTTGTGCC 180
DB 1646 GGGCTGAAGCGTACGGCCCTCGCTAAAGCGGGACGCTACCGCCCAACACCATCTTGTGCC 1705
QY 181 CTTTCCCATCTTCATGAAGCCCAATGATGGTGCTACA 216
DB 1706 CTTTCCCATCTTCATGAAGCCCAATGATGGTGCTACA 1741
LOCUS AK074975 2227 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90434 fis, clone NT2RP3003614.
ACCESSION AK074975
VERSION AK074975.1 GI:22760768
KEYWORDS cDNA capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1
AUTHORS
  Isoqali,T., Ota,J., Nishikawa,T., Hayashi,K., Otsuki,T.,
  Sugiyama,T., Suzuki,Y., Nadei,K., Sugano,S., Ishii,S.,
  Kawai,H.O.Y., Saito,K., Yamamoto,Y., Wakamatsu,A., Nakamura,Y.,
  Kojima,S., Nagahara,K., Masuno,Y., Ota,T., Okano,K., Yoshikawa,Y.,
  Aotsuka,S., Sasaki,N., Pattori,A., Okumura,K., Iwayanagi,I. and
  Ninomiya,K.
  NEDO human cDNA sequencing project
  Unpublished
  2 (bases 1 to 2227)
  Isoqali,T. and Otsuki,T.
  Direct Submission
  Submitted (25-MAR-2002) Takao Isoqali, Helix Research Institute,
  Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
  (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
  NEDO human cDNA sequencing project supported by Ministry of
  Economy, Trade and Industry of Japan; cDNA full insert sequencing:
  Research Association for Biotechnology; cDNA library construction:
  Institute of Medical Science, University of Tokyo, Laboratory of
  Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
  sequencing and clone selection; Helix Research Institute (supported
  by Japan Key Technology Center etc.).
FEATURES
  source
    1..2227
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /cclone="NT2RP3003614"
      /cell_line="NT2"
      /cell_type="teratocarcinoma"
      /clone_lib="NT2RP3"
      /note="cloning vector: pWEL8SFL3-mRNA from NT2 neuronal
      precursor cells after 2-weeks retinoic acid (RA)

```



```

CDS
neuronal precursor cells."
136..1426
/ncbi="unlabeled protein product"
/codon_start=1
/protein_id="RAE56136.1"
/db_xref="gi:14042171"
/transluc="MSVNSEKSYDQWELKRMEDLRASSLYVAFSTCYKVL
GCEKRGCKKTCIASREPKCWKEGACSHLSFNKSHITPECDIERNTPEACCHN
SCEVUNGSSSPETTSSTAQEGYSRGMQDKKHLDSPOSINLGAISSENFQ
DKRGVRSYKAGHQCEVTLIAVLAFYMGAVFSGITVYCVDRRDNVAVCR
KEKEIHSRSGMSVTLKSLGFCQTSKDKPFAITPLKHNGKLA:FGNTAKMLIK
ADQHLDTALSTPESTPTLOOKRKPSRGSREMERONINACTIKDMPKSEVPTD
JLBRSPHISPVVLPITQGYCHEYVDPKMSFVACMA:EDQATLEYKIKHLS
SKSPRGYNLVNLSLSPKVPORASLGPGRASISQGLSKRIEMHSSSYNDYAR
SIPTSLTSHCATLKLNNTSSNSHSNOSRGNDPPFPAPQVDSVQVHSQP
SQQAVTSQPSLAINSLTREGKRTSKDPVPEKPSFAPLSTSMKPNDACT"
BASE COUNT 579 a 713 c 581 g 420 t
ORIGIN
Query Match 100.0%; Score 216; DB 9; Length 2293;
Best Local Similarity 100.0%; Pred. No. 8.4e-42;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCCGCCGCCGCCGAGGGTGGACTCCATCCAGGTGCACAGCTCCGAGCATCTGCG 60
DB 1608 CCGCCGCCGCCGCCGAGGGTGGACTCCATCCAGGTGCACAGCTCCGAGCATCTGCG 1667
QY 61 CAGGCGGTGACTGTCTGAGGAGCAGCCAGCCCTCAAGCTCAACTCACTGACAGGTCG 120
DB 1668 CAGGCGGTGACTGTCTGAGGAGCAGCCAGCCCTCAAGCTCAACTCACTGACAGGTCG 1727
QY 121 GGGCTGAAGCTAGCGCTCGCTAAAGCGGAGGTAAAGCCCAACCACTGCTTGGCTCG 180
DB 1728 GGGCTGAAGCTAGCGCTCGCTAAAGCGGAGGTAAAGCCCAACCACTGCTTGGCTCG 1787
QY 181 CTTCCACATCCATGAAGCCCAATGATCGGTGTACA 216
DB 1788 CTTCCACATCCATGAAGCCCAATGATCGGTGTACA 1823

RESULT #
BD159853
LOCUS
DEFINITION
Homo sapiens cDNA FLJ14748 fis. clone NT2RP3002869, highly similar
to Mus musculus semaphorin 7A mRNA.
ACCESSION
BD159853
VERSION
BD159853.1 GI:27865611
KEYWORDS
JP 2002191363-A/14696.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Orai.T., Isoqai.T., Nishikawa.T., Sugiyama.T., Saito.K., Yamamoto.J.,
Zahit.S., Sugiyama.T., Wakamatsu.A., Nishikawa.T. and Otsuki.T.
TITLE
Primer for synthesizing full-length cDNA and use thereof
JOURNAL
Patent: JP 2002191363-A 14696 69-71-2002;
HEIIX RESEARCH INSTITUTE
COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/14696
PC 28-JUL-2002
PF 28-JUL-2006 JP 20020280990
PI TOSHIO OTA,TAKA ISOQA,TETSUO NISHIKAWA,KOJI HAYASHI,KAGURU
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/89,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N/21,C12N5/PC
10,
PC C12P21/32,C12Q1/66/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS
(109)..(1830).
Location/Qualifiers

```

```

SOURCE
1..2136
/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="taxon:9606"
BASE COUNT 579 a 739 c 577 g 411 t
ORIGIN
Query Match 100.0%; Score 216; DB 6; Length 2306;
Best Local Similarity 100.0%; Pred. No. 8.3e-42;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCCGCCGCCGCCGAGGGTGGACTCCATCCAGGTGCACAGCTCCGAGCATCTGCG 60
DB 1615 CCGCCGCCGCCGCCGAGGGTGGACTCCATCCAGGTGCACAGCTCCGAGCATCTGCG 1674
QY 61 CAGGCGGTGACTGTCTGAGGAGCAGCCAGCCCTCAAGCTCAACTCACTGACAGGTCG 120
DB 1675 CAGGCGGTGACTGTCTGAGGAGCAGCCAGCCCTCAAGCTCAACTCACTGACAGGTCG 1734
QY 121 GGGCTGAAGCTAGCGCTCGCTAAAGCGGAGGTAAAGCCCAACCACTGCTTGGCTCG 180
DB 1735 GGGCTGAAGCTAGCGCTCGCTAAAGCGGAGGTAAAGCCCAACCACTGCTTGGCTCG 1794
QY 181 CTTCCACATCCATGAAGCCCAATGATCGGTGTACA 216
DB 1795 CTTCCACATCCATGAAGCCCAATGATCGGTGTACA 1830

RESULT #
AK027654
LOCUS
DEFINITION
Homo sapiens cDNA FLJ14748 fis. clone NT2RP3002869, highly similar
to Mus musculus semaphorin 7A mRNA.
ACCESSION
AK027654
VERSION
AK027654.1 GI:14042491
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Isoqai.T., Ota.T., Hayashi.K., Sugiyama.T., Otsuki.T., Suzuki.Y.,
Nishikawa.T., Nishikawa.T., Sugano.S., Shikatori.A., Sudo.H.,
Wakamatsu.M., Hosogi.T., Kaku.Y., Kodaira.H., Kondo.H.,
Sugawara.M., Takahashi.M., Chiba.Y., Ishida.S., Murakawa.K.,
Orai.Y., Takiguchi.S., Katanabe.S., Kimura.K., Murakami.K.,
Ishii.S., Kawai.Y., Saito.K., Yamamoto.J., Wakamatsu.A.,
Nakamura.Y., Nishihara.K., Masuno.Y., Ninomiya.K. and Iwayanagi.T.
TITLE
NEBO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2306)
AUTHORS
Isoqai.T. and Otsuki.T.
TITLE
Direct Submission
JOURNAL
Submitted (10-MAY-2001) Takao Isoqai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
1..2306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3002869"
/cell_line="N2"
/cell_type="teratocarcinoma"
/clone_lib="N2RP3"
/note="cloning vector: pME18SPL3-mRNA from NT2 neuronal

```





```
/cds_lib="N12584"
/note="cloning vector: pUC19(+) M13A from mind-loci RT;
national precursor: cbl.1."
89, 12184
/note="unnamed protein product."
/codon_start=1
/protein_id="RAB55156.1"
/db_xref="GI:1434223"
/translation="MIDTASVHTGKFKQKSPDSTWTFPEERVKRPDPCAGSSSL
ERYATNFFPDILNFIKIHIMDEAVPSTNREMLTVMVYATIAVITAAGPYO
NIIVFLSEKIIILKFLRIGNSGFLNDSILFLEMVYNSKSYSEVEDKRMGMQ
LDRASSLVAFSTCIKVLQGRCHGCKKICASRDPCGWLKEGACSHSPNS
RLTFQDIERGTDLGCHKSFVALNGHSSLLPSTTSDIAEGVESKGMCDMK
HLLDPSDTPDLCVSSNHDKKGVRESYKGDQJVPVILLATAVILAPYGVAF
SLITYVCHRRKQDAVVRKEKELTHRRGSMSSVTKLSGFGGCTGSKPKFENIL
TFLMNGKATDNGTAKMLKADQHLDLTALPTPESTPILQKKRPSGRSEWENO
NINAKTDMPGSPVPTPLRASPSRIPSVVLPITQGGQHEVVDQPKNSYA
QVALPDCAATLKYKIKELHSSKPNHGNVNLDSLPKVPQREASLGGPGLSLG
TGLSKLEMHSSSYGVQDKRSYPNLTTRSHQATTEKRNNTNSNSHLSKNGSFGS
GNPPAPORVDSIOVHSSQPSGQAVTVSRDPSLNAYNSLRSGLKATPSLKPDVPRK
PSFAPLSMKPNDACI"
BASE COUNT 804 a 886 c 714 g 637 t
ORIGIN
```

```
Query Match 100.0%; Score 216; DB 9; Length 3041;
Best Local Similarity 100.0%; Pred. No. 8e-42;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGCCGCCGCCGAGAGGGTGGACTCCATCCAGGTCCAGAGTCCAGGCACTGCGG 60
DB 1966 CGCCGCCGCCGCCGAGAGGGTGGACTCCATCCAGGTCCAGAGTCCAGGCACTGCGG 2025

QY 61 CAGGCGGTGACTGTCTCGAGGAGCCAGCCAGCTCAGGCTACAGCTACAGCTAGCTG 120
DB 2026 CAGGCGGTGACTGTCTCGAGGAGCCAGCCAGCTCAGGCTACAGCTACAGCTAGCTG 2085

QY 121 GGCTCAAGGTAGCCGCTCGCTTAAAGCCGAGGACACCGCCCAAGAGTCTTCTGCTCC 180
DB 2086 GGCTCAAGGTAGCCGCTCGCTTAAAGCCGAGGACACCGCCCAAGAGTCTTCTGCTCC 2145

QY 181 CTTCCACATCCATGAAGCCCATATGCGGTGACA 216
DB 2146 CTTCCACATCCATGAAGCCCATATGCGGTGACA 2181
```

```
RESULT 22
AX026741
LOCUS AX026741 3093 bp DNA linear PAT 16-SEP-2003
DEFINITION Sequence 1 from Patent WO0043252.
ACCESSION AX026741
VERSION AX026741.1 GI:10187866
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Klostermann, A. and Bahl, C.
TITLE Human semaphorin 6A (sema6A), a gene involved in neuronal
development and regeneration mechanisms during apoptosis, and its
use as a potential drug target.
JOURNAL Patent: WO 0031252-A 1 02-JUN-2000;
KLOSTERMANN ANDREAS (DE) ; MAX PLANCK GESELLSCHAFT (DE) ; RHEIN
CHRISTIAN (DE)
FEATURES
source
Location/Qualifiers
1..3093
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3093
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09107.1"
```

CDS

```
/db_xref="GI:10187867"
/translation="MRSEALLIYFLLHFAGAGFPEDSRPISISHGNTYKQVPVFGH
KPNVITQRELDIOMINIMNGTIIAARCHIYTVDLTSHTEIYCSKKLTKWRQRA
DVTIRKGGKHKDECHNEIKVLLKKNDDALFVCGTNAFNPCRNKYMOTLEPFGDEFS
GMARCPYDAKHANVALEADGKIY SATVTFDAIDAVIYKSLGESPTLRTVKHDSKWLK
RPVCEADVYGDYIYFEPREIAVEYNTGKVPFVPAQVCKNDKMSORVLEKQWTSF
LKAFACSVPGSHFVFNILQAVTDVIRINGROVVLATFSTPYSNIRGSAYCAVMDLD
IASVFGRKPKSKSPSTTPVDERVAPKPRGCCAGSSILRYATSNFEPDDTLNFI
KTHPDMEAVPSIFNRPWELRTMVRRLTIADVTAAAGPYQHTVYVLSKGIILKE
IAIIGNSGFLNDSILFLEMVYNSKSYSEVEDKRMGMQDRASSLVAFSTCIK
KVLPGCHRGCKKICASRDPCGWLKEGACSHSPNSRLTTEQDIERTGNTDGLV
DCHNSVALNGHSSLLPSTTSDIAEGVESRGMCDMKHLLDPSDTPDLCVSS
HNQDCKGVRESYKAGHDLVPTLLAVILAFVMAVFGSIYVYCVCHRRKQDA
VYORKKELTHRRGSMSSVTKLSGFGGCTGSKPKFENILPLMNGKATPGNTAK
MLIKADQHLDLTALPTPESTPILQKKRPSGRSEWERNQNLINACTKDMPPGSPV
IPDLDLRASPSRIPSVVLPITQGGQHEVVDQPKNSYQTLGSKLEMHSSSYGV
EHLSSKPNHGNVNLDSLPKVPQREASLGGPGLSLGDKRSYPNLTTRSHQATTE
DYKRSYPNLTTRSHQATTEKRNNTNSNSHLSKNGSFGS
SSQPSGQAVTVSRDPSLNAYNSLRSGLKATPSLKPDVPRKPSFAPLSMKPNDACI"
BASE COUNT 813 a 855 c 765 g 660 t
ORIGIN
```

```
Query Match 100.0%; Score 216; DB 6; Length 3093;
Best Local Similarity 100.0%; Pred. No. 8e-42;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGCCGCCGCCGAGAGGGTGGACTCCATCCAGGTCCAGAGTCCAGGCACTGCGG 60
DB 2875 CGCCGCCGCCGCCGAGAGGGTGGACTCCATCCAGGTCCAGAGTCCAGGCACTGCGG 2934

QY 61 CAGGCGGTGACTGTCTCGAGGAGCCAGCTCAGGCTACAGCTACAGCTAGCTG 120
DB 2935 CAGGCGGTGACTGTCTCGAGGAGCCAGCTCAGGCTACAGCTACAGCTAGCTG 2994

QY 121 GGCTCAAGGTAGCCGCTCGCTTAAAGCCGAGGACACCGCCCAAGAGTCTTCTGCTCC 180
DB 2995 GGCTCAAGGTAGCCGCTCGCTTAAAGCCGAGGACACCGCCCAAGAGTCTTCTGCTCC 3054

QY 181 CTTCCACATCCATGAAGCCCATATGCGGTGACA 216
DB 3055 CTTCCACATCCATGAAGCCCATATGCGGTGACA 3090
```

```
RESULT 13
AX099520
LOCUS AX099520 3550 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 160 from Patent WO0119988.
ACCESSION AX099520
VERSION AX099520.1 GI:13538594
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Jacobs, K., McCoy, J.M., Lavallie, E.R., Collins-Kacie, L.A., Evans, C.,
Merberg, D., Treacy, M., Bowman, M.R., Spaulding, V. and Acostino, M.J.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patent: WO 0119988-A 160 22-MAR-2001;
Genetics Institute, Inc. (US)
FEATURES
source
Location/Qualifiers
1..3550
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
957 a 994 c 856 g 742 t 1 others
BASE COUNT 957 a 994 c 856 g 742 t
ORIGIN
```

CDS













```
Db 5647 TCAGACC 5653
: : :
: : :

RESULT 11
US-09-764-891-810
: Sequence 810, Application US/09-764-891
: Publication No. US20030027780A1
: GENERAL INFORMATION:
: APPLICANT: ROSEN et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antipodides
: FILE REFERENCE: PC006
: CURRENT APPLICATION NUMBER: US/09-764-891
: CURRENT FILING DATE: 2002-01-17
: Prior application data removed - consult PAAM or file wrapper
: NUMBER OF SEQ ID NOS: 10231
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 810
: LENGTH: 662
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (321)
: OTHER INFORMATION: n equals a.t.g. or c
US-09-764-891-810

Query Match 24.7% Score 53.4; Db 11; Length 662;
Best Local Similarity 53.8%; Pred. No. 3.5e-07;
Matches 81; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 74 TCTGAGGAGCCAGCCCTCAAGCCCTAGCACTGACAGGTGGGGCTGAGCCGTA 133
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 28 TCCACAGACAGCAGCAGCAGCAGTAATGGCAGCTCTCCAGCAGGACAAAGAGCA 87
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 134 CGCCCTCGTAAAGCGGGAGTACCCCGCAACCATGCTTGGTCCCTTTGACATCA 153
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86 CGCGCTTAAAGACTGAGTGGCACCAGGCTCTCTGTTCTCAAGACCTATCG 147
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 194 TGAAGCC 203
||| |||
DB 146 TCAGACC 154
||| |||

RESULT 12
US-10-027-632-166826/c
: Sequence 166826, Application US/10027632
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: Polymorphisms in the Human Genome
: FILE REFERENCE: 108627.124
: CURRENT APPLICATION NUMBER: US/10/027.632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,036
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,062
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 166826
: LENGTH: 765
: TYPE: DNA

Query Match 17.3% Score 37.4; Db 13; Length 765;
Best Local Similarity 55.9%; Pred. No. 0.036; Mismatches 56; Indels 0; Gaps 0;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 33 CCAGGTGCAGAGCTCCAGGCGATGTGGCAGGCGGTGACTGTCTGAGGACGCCAGCC 92
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 268 CTTGCTCCAGGCGGCGCCAGCATCTGGCGGTGGCAGACACTGGCGTGGCTGGCGGACC 209
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 CAACGCGCTACACACACTGACAAAGTCCAGGCTGAGGCTAGCGGCTGCTTAAAGCGGCA 152
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 238 AGAGGGCGTTTCCACAGGTACAAAGTGGCGGTGGAGGCCAGGACCCCTGGACAGCC 149
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 153 CGTACCC 159
||| |||
DB 148 CCATCCC 142
||| |||

RESULT 13
US-09-554-000-7/c
: Sequence 7, Application US/09554000
: Patent No. US2002015364A1
: GENERAL INFORMATION:
: APPLICANT: Islen, Roger Y.
: APPLICANT: Miyawaki, Atsushi
: TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
: TITLE OF INVENTION: DETECTION OF ANALYTES
: FILE REFERENCE: 07257/042001
: CURRENT APPLICATION NUMBER: US/09/554,000
: CURRENT FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: 08/818,252
: PRIOR FILING DATE: 1997-03-14
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 1971
: TYPE: DNA
: ORGANISM: Aequorea victoria
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (0)....(1958)
US-09-554-000-7

Query Match 16.4% Score 35.4; Db 10; Length 1971;
Best Local Similarity 61.3%; Pred. No. 0.17;
Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 2 CGCCCGCCGCGCGCGCGAGGAGGTGGACTCCATCCAGGTGCAGAGCTCCAGGCACTCTGCC 61
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 112 CGCGCTCCAGTCCAGCAGGATGGGACACCCCGGTGAACAGCTCTGGCGCTGTCTCA 53
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 AGGCGGTGACTGTCTGAGGCGAGGCGGAGCCTCA 94
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 92 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-10-128-714-2578
: Sequence 2578, Application US/10:28714
: Publication No. US2003011901A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Ho
: APPLICANT: HS, Wenqi
: APPLICANT: Tiskoff, Daniel
: APPLICANT: Zarudio, Carlos
: APPLICANT: Ershkin, Alexey M
: APPLICANT: Lemieux, Sebastien M
: TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus
: TITLE OF INVENTION: Methods of Use
: FILE REFERENCE: 10182-018-999
: CURRENT APPLICATION NUMBER: US/10/128,714
```





GenCore version 5.1.0  
Copyright (c) 1993 - 2003 CorpGen Inc.

OM protein - protein search, using sw model

Run on: September 10, 2003, 16:32:13 : Search time 445 seconds  
(without alignments)  
24.461 Million cell updates/sec

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAPQVDSIQVSSQPSG.....PTKSPAPSTISMKPNDACI 72

Scoring table: B:OSCM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Published\_Applications\_AA:\*

1: /cqn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*

2: /cqn2\_6/prodata/2/pubpaa/PCR\_NEW\_PUB.pep.\*

3: /cqn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cqn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*

5: /cqn2\_6/prodata/2/pubpaa/US06\_PUB.pep.\*

6: /cqn2\_6/prodata/2/pubpaa/PCRUS\_PUBCOMB.pep.\*

7: /cqn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cqn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*

9: /cqn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cqn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cqn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cqn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cqn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cqn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cqn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cqn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*

17: /cqn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*

18: /cqn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	376	100.0	1047	11	US-09-957-187-85 Sequence 85, Appl
2	86	22.9	973	12	US-09-864-245-3 Sequence 3, Appl
3	80.5	21.4	374	12	US-10-233-431A-20 Sequence 20, Appl
4	80.5	21.4	564	12	US-10-233-431A-3 Sequence 8, Appl
5	74	19.7	629	15	US-10-213-595-36 Sequence 36, Appl
6	74	19.7	884	15	US-10-303-683-22 Sequence 21, Appl
7	74	19.7	884	15	US-09-291-417-18 Sequence 18, Appl
8	74	19.7	884	15	US-09-303-683-20 Sequence 20, Appl
9	73.5	19.5	2665	9	US-09-864-761-34248 Sequence 34248, A
10	73.5	19.5	3664	15	US-10-177-293-424 Sequence 423, Appl
11	73	19.4	200	12	US-10-259-165-300 Sequence 300, Appl
12	73	19.4	4019	10	US-09-738-973-425 Sequence 423, Appl
13	73	19.4	4019	10	US-09-854-133-425 Sequence 423, Appl
14	73	19.4	4019	15	US-10-144-649A-425 Sequence 423, Appl
15	71.5	19.0	223	9	US-09-864-761-33417 Sequence 33417, A

Sequence 1116, Ap  
Sequence 2, Appl  
Sequence 152, App  
Sequence 22, Appl  
Sequence 294, App  
Sequence 11, Appl  
Sequence 1631, Ap  
Sequence 122, App  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 5, Appl  
Sequence 1, Appl  
Sequence 6, Appl  
Sequence 218, App  
Sequence 1068, Ap  
Sequence 1068, Ap  
Sequence 1068, Ap  
Sequence 24, Appl  
Sequence 170, App  
Sequence 44230, A  
Sequence 527, App  
Sequence 77, Appl  
Sequence 166, App  
Sequence 140, App  
Sequence 142, App  
Sequence 38306, A  
Sequence 80, Appl  
Sequence 38321, A  
Sequence 2, Appl

9 US-09-925-301-1116  
9 US-09-972-086-2  
15 US-10-100-957A-152  
15 US-10-100-957A-22  
625 US-09-801-368-294  
225 US-10-145-396-11  
153 US-09-925-300-1631  
632 US-09-934-455-122  
15 US-10-218-957-4  
612 US-10-218-957-2  
1149 US-09-969-528-5  
258 US-05-922-543-1  
274 US-10-188-702A-6  
465 US-10-153-668-218  
5179 US-09-922-217-1069  
9 US-09-833-253-1068  
5179 US-10-025-360-1068  
503 US-10-078-547-2  
507 US-10-078-547-24  
200 US-10-062-254-170  
448 US-09-864-761-44230  
479 US-09-925-302-527  
144 US-10-263-828-77  
187 US-10-062-254-168  
261 US-10-062-254-140  
261 US-10-062-254-142  
276 US-09-864-761-38306  
463 US-10-039-180-80  
473 US-09-864-761-38321  
1142 US-09-899-651-2

#### ALIGNMENTS

RESULT 1  
US-09-957-187-85  
Publication 85, Application US/09957187  
Publication No. US2003005414A1  
GENERAL INFORMATION:  
APPLICANT: Shimkets, Richard A.  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY  
FILE REFERENCE: 15966-540 CIP  
CURRENT APPLICATION NUMBER: US/09/957,187  
CURRENT FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/123,667  
PRIOR FILING DATE: 1999-04-09  
PRIOR APPLICATION NUMBER: 09/520,761  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/234,082  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 60/233,798  
PRIOR FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/174,485  
PRIOR FILING DATE: 2000-01-04  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 85  
LENGTH: 1047  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-957-187-85

Query Match 100.0% Score 376; DB 11; Length 1047;  
Best Local Similarity 100.0% Pred. No. 1.9e-29;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVSSQPSGQAVTVSRQPSLNNYSITRSGLKRTPSLKPDVPPKPSFAP 60  
DB 976 PPPAPQVDSIQVSSQPSGQAVTVSRQPSLNNYSITRSGLKRTPSLKPDVPPKPSFAP 1035  
CY 61 LSTSMKPNDACT 72

1036 LSTSMKPNDACT 1047

```

RESULT 2
US-09-843-245-3
US-09-843-245-3, Application: US/09843245
? ? Patent No. US2002016472A1
? ? GENERAL INFORMATION:
? ? APPLICANT: McPherson, Peter S.
? ? APPLICANT: Ramjaun, Antoine Rachid
? ? TITLE OF INVENTION: REGULATION OF
? ? TITLE OF INVENTION: INTERACTION B
? ? TITLE OF INVENTION: AND THE GERM
? ? FILE REFERENCE: 9555_116G501
? ? CURRENT APPLICATION NUMBER: US/09/
? ? CURRENT FILING DATE: 2001-04-25
? ? NUMBER OF SEQ ID NOS: 13
? ? SOFTWARE: PatentIn Ver. 2.1
? ? SEQ ID NO 3
? ? LENGTH: 873
? ? TYPE: PR
? ? ORGANISM: Rattus sp.
US-09-843-245-3

```

```

RESULT 3
US-10-239-431A-20
; Sequence 20, Application US/10139431A
; Publication No. US20030170725A;
; GENERAL INFORMATION:
; APPLICANT: FRADELIZE, JULIE
; APPLICANT: FRIEDERICH, EVELYNE
; APPLICANT: GOLSTEYN, ROY M.
; APPLICANT: LOUVAUD, DANIEL
; APPLICANT: NOIREAUX, VINCENT
; APPLICANT: SYKES, CECILE
; TITLE OF INVENTION: PEPTIDE SEQUENCES COMPRISING ONE OR MORE UNITS BINDING
; TO PROTEINS OF THE ENA/VASP FAMILY, AND THEIR USES
; FILE REFERENCE: 0509-1032
; CURRENT APPLICATION NUMBER: US/10/239,431A
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PC1/FR01/00843
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: FR 00/03637
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentia Ver. 2.1
; SEQ ID NO 20
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Murine sp.
US-10-239-431A-20

```

```

Query Match      21.4%   Score 80.5; DB 12; Length 374;
Best Local Similarity 32.8%; Pred. No. 2.6;
Matches 22; Conservative 11; Mismatches 25; Indels 5; Gaps 4

QY 1 PPAFQVDSDVHVSQSSGQAVTVSRQPSLNAYNSLTRAQKRTVSLKPPVPPKPSAP 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 208 PPFQFQAKPQVQJH-VQQAAP-HVQQAAP-VSSANTQPRGLSQAT-----PAPKAP 258
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

: RESULT 4
: RS-10-239-431A-8
: Sequence 8, Application US/10239431A
: Publication No. J620030-70726A1
: GENERAL INFORMATION:
: APPLICANT: FRADELIZE, JULIE
: APPLICANT: FRIEDERICH, EVELYNE
: APPLICANT: GOLSTEYN, ROY M.
: APPLICANT: LOIVARD, DANIEL
: APPLICANT: NOIRPAUX, VINCENT
: APPLICANT: SYKES, CECILIE
: TITLE OF INVENTION: PEPTIDE SEQUENCES COMP
: TITLE OF INVENTION: TO PROTEINS OF THE EN
: FILE REFERENCE: 0508-1032
: CURRENT APPLICATION NUMBER: US/0/239,431A
: CURRENT FILING DATE: 2002-09-23
: PRIOR APPLICATION NUMBER: PCT/FR01/00843
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: FR 00/63637
: PRIOR FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patent: Ver. 2.1
: SEQ ID NO 8
: LENGTH: 564
: TYPE: PRT
: ORGANISM: Murine sp.
: RS-10-239-431A-8

```

```

1 RESULT 5
2 US-10-213-990-36
3 ? Sequence 36, Application US10213990
4 ? Publication No. US2003008259SA1
5 ? GENERAL INFORMATION:
6 ? APPLICANT: Jiang, Bo
7 ? APPLICANT: Bussey, Howard
8 ? APPLICANT: Storms, Reg
9 ? APPLICANT: Roemer, Terry
10 ? TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUNIGATUS ENCODING INDUSTRIAL
11 ? TITLE OF INVENTION: ENZYMES AND METHODS OF USE
12 ? FILE REFERENCE: 20182-013-999
13 ? CURRENT APPLICATION NUMBER: US/10/213,990
14 ? CURRENT FILING DATE: 2002-08-05
15 ? NUMBER OF SEQ ID NOS: 72
16 ? SOFTWARE: FastSeq for Windows Version 4.0
17 ? SEQ ID NO 36
18 ? LENGTH: 628
19 ? TYPE: PRT
20 ? ORGANISM: Aspergillus
21 US-10-213-990-36

```

```
QY 6 QVDSQVSSQSGAVTVSRQPSLNAYNSL--TRSGLKRTPS 48
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 QVDSNPINFGAPAGLEVTSRDPVSNQWPSRAQRQGVASDQ 222
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-10-303-683-21
; Sequence 21, Application US/10-04583
; Publication No. US20030113782A1
; GENERAL INFORMATION:
; APPLICANT: EXCELIXIS, INC.
; TITLE OF INVENTION: MAP4KS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE
; FILE REFERENCE: EXG2-126C
; CURRENT APPLICATION NUMBER: US/10/303,583
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/333,378
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; TYPE: PRT
; LENGTH: 884
; ORGANISM: Homo sapiens
US-10-303-683-21

Query Match 19.7%; Score 74; DB 15; Length 884;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

QY 1 PPPAPQVDSI----QVHSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPDVPEKP 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 PPPLPKPKSIPTQEMHSTEDENOG-TTKRCP-----MSGSPAKPSQVPPRPPPP 471
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 SFAP 60
Db 472 RLPP 475

RESULT 7
US-09-291-417-18
; Sequence 18, Application US/09291417A
; Publication No. US2003005023CA1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARINEZ, RICARDO
; APPLICANT: WEXEL, DAVID
; TITLE OF INVENTION: STR20-RELATED PROTEIN KINASES
; FILE REFERENCE: 249/300
; CURRENT APPLICATION NUMBER: US/09/291,417A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: US 60/081,784
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; TYPE: PRT
; LENGTH: 894
; ORGANISM: Mammalian (Human) KES2
US-09-291-417-18

Query Match 19.7%; Score 74; DB 15; Length 894;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

QY 1 PPPAPQVDSI----QVHSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPDVPEKP 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 PPPLPKPKSIPTQEMHSTEDENOG-TTKRCP-----MSGSPAKPSQVPPRPPPP 481
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 SFAP 60
Db 482 RLPP 485
```

```
RESULT 8
US-10-303-683-20
; Sequence 20, Application US/10303683
; Publication No. US20030113782A1
; GENERAL INFORMATION:
; APPLICANT: EXCELIXIS, INC.
; TITLE OF INVENTION: MAP4KS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE
; FILE REFERENCE: EXG2-126C
; CURRENT APPLICATION NUMBER: US/10/303,683
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/333,378
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 894
; ORGANISM: Homo sapiens
US-10-303-683-20

Query Match 19.7%; Score 74; DB 15; Length 894;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

QY 1 PPPAPQVDSI----QVHSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPDVPEKP 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 PPPLPKPKSIPTQEMHSTEDENOG-TTKRCP-----MSGSPAKPSQVPPRPPPP 481
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 SFAP 60
Db 482 RLPP 485

RESULT 9
US-09-864-761-34248
; Sequence 34248, Application: US/09864761
; Patent No. US20020348763A1
; GENERAL INFORMATION:
; APPLICANT: Penn. Shartor G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```

1

```

Best Local Similarity 30.4%; Pred. No. 7.3;
Matches 21; Conservative 11; Mismatches 31; Indels 4; Gaps 3;

QY 2 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
Db 129 PAVQAVAVQCO:SAISSPAPRISTGKAPRVYHETKAGTGRKSPATTFSTPPASER 186
QY 58 FAPLSTSMK 66
Db 189 AATCASSER 157

RESULT 12
US-09-738-973-425
: Sequence 425, Application US/0373553
: Patent No. US2002010563A1
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Fling, Steven P.
: APPLICANT: Monamath, Raodoh
: APPLICANT: Algate, Paul A.
: APPLICANT: Secrist, Heather
: APPLICANT: Indrias, Carol Yoseph
: APPLICANT: Benson, Darin R.
: APPLICANT: Elliot, Mark
: APPLICANT: Mannion, Jane
: APPLICANT: Kalos, Michael D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.475C9
: CURRENT APPLICATION NUMBER: 03/09/738,973
: CURRENT FILING DATE: 2000-12-14
: NUMBER OF SEQ ID NOS: 587
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 425
: LENGTH: 4019
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-738-973-425

Query Match 19.4%; Score 73; DB 16; Length 4019;
Best Local Similarity 32.5%; Pred. No. 2.3e+02;
Matches 27; Conservative 11; Mismatches 27; Indels 14; Gaps 6;

QY 1 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
Db 910 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
QY 47 --SLKPDVFKPSFAH:STSMKP 67
Db 968 SFSRNSAAPVENCITLSSVSRP 990

RESULT 13
US-09-854-133-425
: Sequence 425, Application US/09854133
: Patent No. US20020183499A1
: GENERAL INFORMATION:
: APPLICANT: Lodes, Michael J.
: APPLICANT: Monamath, Raodoh
: APPLICANT: Henderson, Robert A.
: APPLICANT: Benson, Darin R.
: APPLICANT: Secrist, Heather
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.475C11
: CURRENT APPLICATION NUMBER: 05/09/854,133
: CURRENT FILING DATE: 2001-05-11
: NUMBER OF SEQ ID NOS: 735
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 425

```

```

: LENGTH: 4019
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-854-133-425

Query Match 19.4%; Score 73; DB 10; Length 4019;
Best Local Similarity 32.5%; Pred. No. 2.3e+02;
Matches 27; Conservative 11; Mismatches 27; Indels 18; Gaps 6;

QY 1 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
Db 910 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
QY 47 --SLKPDVFKPSFAH:STSMKP 67
Db 968 SFSRNSAAPVENCITLSSVSRP 990

RESULT 14
US-10-144-649A-425
: Sequence 425, Application US/0144649A
: Publication No. US2003011859A1
: GENERAL INFORMATION:
: APPLICANT: Lodes, Michael J.
: APPLICANT: Wang, Tongtong
: APPLICANT: Fan, Liqun
: APPLICANT: Algate, Paul A.
: APPLICANT: McNeill, Patricia D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.475C11
: CURRENT APPLICATION NUMBER: US/10/-44,649A
: CURRENT FILING DATE: 2002-08-21
: NUMBER OF SEQ ID NOS: 749
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 425
: LENGTH: 4019
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-144-649A-425

Query Match 19.4%; Score 73; DB 15; Length 4019;
Best Local Similarity 32.5%; Pred. No. 2.3e+02;
Matches 27; Conservative 11; Mismatches 27; Indels 18; Gaps 6;

QY 1 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
Db 910 PPAPOQVUS:QVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTKTP-----46
QY 47 --SLKPDVFKPSFAH:STSMKP 67
Db 968 SFSRNSAAPVENCITLSSVSRP 990

RESULT 15
US-09-864-751-33417
: Sequence 33417, Application US/09864761
: Patent No. US20020048761A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David K.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
: FILE REFERENCE: Aecmica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw mode.

Run on: September 30, 2003, 13:15:08 : Search time 72 seconds  
(without alignments)  
158,727 million bytes input/output

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAFQVRDVCVHSSQPSG.....PPQSFALSTSMKPLNDAT 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 6.5

Searched: 1107663 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq\_emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	376	100.0	72	21	AAV71461 Binding domain of
2	376	100.0	507	22	AAV71461 Human protein seqs
3	376	100.0	562	22	AAV71461 Human polypeptide
4	376	100.0	562	22	AAV71461 Human protein seqs
5	376	100.0	574	22	AAV71461 Human protein seqs
6	376	100.0	599	22	AAV71461 Human protein seqs
7	376	100.0	863	22	AAV71461 Novel human diagno
8	376	100.0	974	19	AAV71461 Human secreted pro
9	376	100.0	975	22	AAV71461 Human G145_1 pro

10	376	100.0	1050	21	AAV71460 Human semaphorin 6
11	163.5	43.5	451	22	AAV71460 Human protein seqs
12	163.5	43.5	464	22	AAV71460 Human protein seqs
13	163.5	43.5	474	24	AAV71460 Human MBDT polyep
14	163.5	43.5	1017	23	AAV71460 CAP4P-2, Inocyte 10
15	163.5	43.5	1032	23	AAV71460 Human semaphorin 1
16	163.5	43.5	1035	23	AAV71460 Human semaphorin 6
17	163.5	43.5	1062	22	AAV71460 Arabin acid sequenc
18	163.5	43.5	1082	23	AAV71460 Human semaphorin-1
19	86	22.4	473	24	AAV71460 Rat germinall centr
20	80.5	21.4	374	22	AAV71460 Murine zyxine frag
21	80.5	21.4	564	22	AAV71460 Murine zyxine frag
22	80.5	21.4	564	23	AAV71460 Murine zyxine frag
23	77.5	20.6	693	23	AAV71460 Human polypeptide
24	77.5	20.6	1111	23	AAV71460 Herbicidal y activ
25	75.5	20.1	735	22	AAV71460 Prosophila melanog
26	75.5	20.1	745	22	AAV71460 Prosophila melanog
27	75	19.9	95	22	AAV71460 Prosophila melanog
28	74.5	19.8	449	22	AAV71460 Prosophila melanog
29	74.5	19.8	1475	22	AAV71460 Prosophila melanog
30	74	19.7	873	22	AAV71460 Human protein seq
31	74	19.7	894	22	AAV71460 Human KHS2 protein
32	74	19.7	894	22	AAV71460 Human KHS2 protein
33	74	19.7	930	22	AAV71460 Human protein seq
34	74	19.7	930	22	AAV71460 Human protein seq
35	73.5	19.5	425	22	AAV71460 Novel human diagno
36	73.5	19.5	960	22	AAV71460 Novel human diagno
37	73.5	19.5	960	22	AAV71460 Novel human diagno
38	73.5	19.5	2665	22	AAV71460 Human liver peptid
39	73.5	19.5	2665	22	AAV71460 Human peptide #965
40	73.5	19.5	2665	22	AAV71460 Peptide #996 encod
41	73.5	19.5	2665	22	AAV71460 Protein #949 encod
42	73.5	19.5	2665	22	AAV71460 Human brain expres
43	73.5	19.5	2665	22	AAV71460 Human bone marrow
44	73.5	19.5	2665	22	AAV71460 Peptide #967 encod
45	73.5	19.5	2665	22	AAV71460 Peptide #987 encod

#### ALIGNMENTS

#### RESULT 1

AAV71461

ID AAV71461 standard; peptide; 72 AA.

XX AAV71461;

XX 04-OCT-2003 (first entry)

XX Binding domain of human semaphorin 6A-1.

XX Human: semaphorin 6A-1; (HSA)SEMA6A-1; neuronal development; apoptosis;  
XX neuronal regeneration; Fna/VASP protein family; immunomodulatory;  
XX gene therapy; diagnostic agent; therapeutic agent; differentiation;  
XX cytoskeletal stabilisation; plasticity.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Binding-site 51..56

XX Note= "Specific binding motif for members of Fna/VASP protein family, especially Evi"

XX WO200031252-A1.

XX 02-JUN-2003.

XX 26-NOV-1993; 99WO-EP09215.

XX 26-NOV-1998; 98EP-G122441.

XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.



PI Behl C, Klostermann A;  
 XX WPI: 2000-400065/34.  
 DR N-PSDB: AAD01234.  
 XX  
 PT Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,  
 PT therapeutic agent, for modulating immune system, in gene therapy or for  
 PT effecting differentiation, cytoskeletal stabilization and/or plasticity  
 PT -  
 XX  
 PS Disclosure: Page 22; 53pp; English.  
 XX  
 CC The present sequence is a binding domain of transmembranes  
 CC human semaphorin 6A-1 ((HSA)SEMA6A-1) which is involved in neuronal  
 CC development and regeneration mechanisms during apoptosis. The binding  
 CC domain shows homology to zyxin protein and selectively binds to members  
 CC of Eno/VASP protein family, especially Evl. (HSA)SEMA6A-1 is a  
 CC member of protein family displaying secreted or transmembrane-based  
 CC repulsive guidance cues critically involved in neuronal development.  
 CC Expression of (HSA)SEMA6A-1 is highest in embryonic brain and  
 CC kidney and moderate in lung. The present sequence is useful as diagnostic  
 CC and therapeutic agents for modulating the immune system, in gene  
 CC therapy, for effecting differentiation, cytoskeletal stabilisation  
 CC and plasticity.  
 XX  
 SQ Sequence 72 AA:  
 Query Match 100.0%; Score 376; DB 21; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-25;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQVDSIQVHSSQSGQAVTVSRQPSLNAYNSLTRSLKTPSLKTPVPRKPSFAP 60  
 DB 1 PPPAPQVDSIQVHSSQSGQAVTVSRQPSLNAYNSLTRSLKTPSLKTPVPRKPSFAP 60  
 QY 61 LSTSMKPNDACT 72  
 DB 61 LSTSMKPNDACT 72  
 RESULT 2  
 AAB92688  
 ID AAB92688 standard; Protein: 507 AA.  
 XX  
 AC AAB92688;  
 XX  
 XX 26-JUN-2001 (first entry)  
 DT Human protein sequence SEQ ID NO:11073.  
 DE  
 DE Human; primer: detection: diagnosis; antisense therapy; gene therapy.  
 KW Homo sapiens.  
 OS  
 XX EP1074417-A2.  
 PN 07-FEB-2001.  
 XX  
 XX 28-JUL-2000; 2000EP-0116126.  
 PF  
 XX 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183765.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI: 2001-318749/34.  
 DR

PI Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 XX Claim 8; SEQ ID 11073; 2537pp + CD ROM; English.  
 PS  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and  
 CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAB13629 to AAB13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 507 AA:  
 Query Match 100.0%; Score 376; DB 22; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQVDSIQVHSSQSGQAVTVSRQPSLNAYNSLTRSLKTPSLKTPVPRKPSFAP 60  
 DB 496 PPPAPQVDSIQVHSSQSGQAVTVSRQPSLNAYNSLTRSLKTPSLKTPVPRKPSFAP 495  
 QY 61 LSTSMKPNDACT 72  
 DB 496 LSTSMKPNDACT 507  
 RESULT 3  
 AAB93444  
 ID AAB93444 standard; Protein: 562 AA.  
 XX  
 AC AAB93444;  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT Human polypeptide. SEQ ID NO: 3088.  
 DE  
 DE Human; full length cDNA; cDNA synthesis; oligo-capping.  
 KW Homo sapiens.  
 OS  
 XX EP110094-A2.  
 PN 05-SEP-2001.  
 XX  
 XX 07-JUL-2000; 2000EP-0114089.  
 PF  
 XX 08-JUL-1999; 99JP-0194486.  
 PR 11-JAN-2000; 2000JP-0118774.  
 PR 02-MAY-2000; 2000JP-0183765.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA  
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koqa H;  
 DR

XX WPI: 2001-524355/58.  
 DR N-PSDS: AAK94365.  
 XX  
 PT 830 Primers useful for synthesizing full length cDNA clones and their  
 use in genetic manipulation.  
 XX  
 PS Claim 8: SEQ ID NO 3088: 1380bp + sequence listing: English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 clones. 830 cDNA molecules encoding a human protein have been  
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
 CC molecules have been determined. Primers for synthesizing the full length  
 CC cDNA are useful for clarifying the function of the protein encoded by  
 CC the cDNA. The full length clones were obtained by construction of full  
 CC length enriched cDNA libraries that were synthesised by the oligo-priming  
 CC method. The primers enable the production of the full length cDNA easily  
 CC without any special methods. The present sequence is a polypeptide  
 CC encoded by a full length human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the claimed  
 CC specification, but was obtained in CD-ROM format directly from EPO.  
 XX  
 SQ Sequence 562 AA:  
 Query Match 100.0%; Score 376; DB 22; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQRVDSTQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60  
 DB 491 PPPAPQRVDSTQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 550  
 QY 61 LSTSMKPNDACT 72  
 DB 55: LSTSMKPNDACT 562  
 RESULT 4  
 AAB94104  
 ID AAB94104 standard: Protein; 562 AA.  
 XX AAB94104:  
 XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:14328.  
 XX Human: primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241699.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isozaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 WPI: 2001-318749/34.  
 XX  
 DR Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the

PI full-length cDNAs -  
 XX Claim 8: SEQ ID 14328, 2517bp - CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence and 3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAB93166 to AAB93628 and  
 CC AAB93633 to AAB94742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAB93629 to AAB93632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 562 AA:  
 Query Match 100.0%; Score 376; DB 22; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQRVDSTQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60  
 DB 491 PPPAPQRVDSTQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 550  
 QY 61 LSTSMKPNDACT 72  
 DB 55: LSTSMKPNDACT 562  
 RESULT 5  
 AAB95317  
 ID AAB95317 standard: Protein; 574 AA.  
 XX AAB95317:  
 XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:17566.  
 XX Human: primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241699.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isozaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI: 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs.

XX

PS Claim 8: SEQ ID 17568: 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC polynucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialized methods. AAB95139 to AAB95142 and

CC AAB95133 to AAB95142 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAB93629 to AAB93632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

SQ Sequence 574 AA:

Query Match 100.0%; Score 376; DB 22; Length 574;

Best Local Similarity 100.0%; Pred. No. 9, 30-34;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSAP 60

DB 503 PPPAPRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSAP 592

QY 61 LSTSMKPNDACT 72

DB 563 LSTSMKPNDACT 574

RESULT 6

AAB95139

XX AAB95139 standard; Protein: 599 AA.

XX

AC AAB95139;

XX

XX

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO: 7154.

XX

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

XX EP1074617-A2.

PN

XX

PD 07-FEB-2001.

XX

XX 28-JUL-2000; 2000EP-0116126.

XX

XX 29-JUL-1999; 99JP-0249036.

PR

XX 27-AUG-1999; 99JP-030253.

PR

XX 11-JAN-2000; 2000JP-0118776.

PR

XX 02-MAY-2000; 2000JP-0133767.

PR

XX 09-JUN-2000; 2000JP-0241899.

XX

FA (HELI-) HELIX RES INCT.

XX

PI ora T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Naga K, Otsuki T;

XX

DB WPI: 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs.

XX

PS Claim 8: SEQ ID 17154: 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialized methods. AAB93629 to AAB93632 and

CC AAB93633 to AAB93642 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAB93629 to AAB93632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

SQ Sequence 699 AA:

Query Match 100.0%; Score 376; DB 22; Length 699;

Best Local Similarity 100.0%; Pred. No. 1, 2e-33;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSAP 60

DB 628 PPPAPRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSAP 587

QY 61 LSTSMKPNDACT 72

DB 686 LSTSMKPNDACT 699

RESULT 7

ARG04056

XX ARG04056 standard; Protein: 863 AA.

XX

AC ARG04056;

XX

XX 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #4057.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

XX W0200175067-A2.

PN

XX 11-OCT-2001.

PD

XX 30-MAR-2001; 2001WO-US08631.

XX















GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2003, 16:21:28 ; Search time 44 seconds

(without alignments)  
30,345 Million cell index/seq

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAPQVDSIQVHSSQSPSG.....PFRKSPAEISSTSMKPKACE 72

Scoring table: BLOSUM62

Gapop 10.0 / Gapext 0.5

Searched: 283308 seqs, 96164682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

PTR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	87	23.1	961	2 A56360	faciogenital dyspl
2	80.5	21.4	1322	2 A50288	myosin heavy chain
3	77.5	20.6	1111	2 T05246	hypothetical prote
4	77	20.5	175	2 T47453	serine/proline-ric
5	76.5	20.3	744	2 E86255	hypothetical prote
6	75.5	20.1	393	2 T33103	fib-1 protein - Ca
7	75.5	20.1	494	2 A42170	zinc finger protei
8	75.5	20.1	457	2 JG5076	myo-associated zin
9	74	19.7	452	2 S22159	imidazoleglycerol
10	73.5	19.5	2882	2 T42717	RNA-binding protei
11	73	19.4	867	2 T41309	hypothetical zinc-
12	72.5	19.3	628	2 S01955	hypothetical prote
13	72.5	19.3	457	2 S84869	probable SRF6 prot
14	72.5	19.3	4357	2 T03455	ALK protein - huma
15	72.5	19.3	5262	2 T03454	ALK protein - huma
16	72	19.1	459	2 A41277	retinoid acid rece
17	71.5	19.0	446	2 A42229	transcription fact
18	71.5	19.0	1522	2 T35471	transcription fact
19	71.5	19.0	2957	2 T33152	hypothetical prote
20	71	18.9	621	2 JG7278	adaptor protein co
21	71	18.9	1150	2 S56775	myo1 protein - stu
22	70.5	18.8	468	2 T46815	hypothetical prote
23	70.5	18.8	1420	2 T33781	probable cytoskele
24	70	18.6	719	2 S62466	probable ATP-depen
25	70	18.6	747	2 S33546	ATP-dependent SMA
26	70	18.6	792	2 T26050	hypothetical prote
27	70	18.6	1012	2 T53172	hypothetical prote
28	70	18.6	1201	2 G85441	unknown protein li
29	69.5	18.5	331	2 B47236	zinc-finger protei

30	59.5	18.5	477	2 A47236	zinc-finger protei
31	49.5	18.5	1125	2 B41206	microtubule-associ
32	69	18.4	3942	2 T42730	Bassoon protein -
33	68.5	18.2	623	2 S48941	regulatory protein
34	68.5	18.2	1106	2 T31742	hypothetical prote
35	68.5	18.2	1189	2 S49915	extensin-like prot
36	68	18.1	423	2 TV000K	transforming prote
37	68	18.1	533	2 A45690	transactivator EBN
38	68	18.1	513	2 T47975	auxin response fac
39	68	18.1	736	2 T25447	hypothetical prote
40	68	18.1	943	2 T40873	probable transcrip
41	68	18.1	1172	2 T00065	hypothetical prote
42	68	18.1	1215	2 T61773	co-repressor prote
43	68	18.1	1223	2 A56068	co-repressor prote
44	68	18.1	1238	2 JG5765	histonol polyphosp
45	67.5	18.0	429	2 JG4965	eik1 protein - mou

#### ALIGNMENTS

##### RESULT 1

A55360

faciogenital dysplasia-associated protein FGJ1 - human

C:Species: Homo sapiens (man)

C:Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 17-Mar-1999

C:Accession: A55360

K:Features: N.G.; Cadle, A.; Lajler, L.J.; Porteous, M.E.M.; Schwartz, C.E.; Stevens

Cell 79, 669-676, 1994

A:Title: Isolation and characterization of the faciogenital dysplasia (Aarskog-Sco

A:Reference number: A55360; MFID:95042764; PMID:7954831

A:Accession: A55360

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-961 <PAS>

A:Cross-references: GB:O11690; NID:G595424; PID:G595425

C:Superfamily: CDC24 homology; p10ckstrin repeat homology

P:573-561/Domain: CDC24 homology <CD24>

Query Match 23.1% Score 87; DB 2; Length: 961;

Best Local Similarity 34.8%; Pred. No. 1.2;

Matches 23; Conservative 5; Mismatches 22; Indels 16; Gaps 2;

QY 2 PPAPQVESIQVHSSQSPSGQAVTVSRGSLNAYNSLTRSGIKRTPSLKPOVPPKPSAPL 61

DB 127 PEGPQRU-----NSDQGFPTETPSQRP-----SPLKRAPGKPPQVPPPSYLOM 170

QY 52 STSMKP 67

DB 171 PRMPPE 176

##### RESULT 2

A59288

myosin heavy chain Myr 8 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000

C:Accession: A59288

R:Patel, K.G.; Liu, C.; Cameron, P.L.; Cameron, R.S.

submitted to GenBank, November 1999

A:Description: Identification of a Novel Mammalian Myosin Class, XVI, in Developing

A:Reference number: A59288

A:Accession: A59288

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1322 <PAI>

A:Cross-references: GB:AF209114; PID:AAF20150.1

A:Experimental source: strain Sprague-Dawley; clone KP4; cell type type 1 astrocyte

C:Superfamily: myosin motor domain homology

F:404-1132/Domain: myosin motor domain homology <MYO>

Query Match 21.4% Score 80.5; DB 2; Length 1322;

Best Local Similarity 38.4%; Pred. No. 7.4;

```

Matches 26: Conservative 7: Mismatches 25: Indels 13: Gaps 4:

QY 3 PAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLRS-GLKPTPSLKRFPVRFKFTAP 60
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 1248 PVPMAVDNLALALASPS-----SRSTGLHVFPSKLSTGQ---PSPRKQFHPKPKAPN 1248
QY 61 --LSTSMKPNQAC 71
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 1299 TPLSASYFAVSAC 1311

RESULT 3
T05646
Hypothetical protein F20D10.310 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
R:Accession: T05646
R:Author: M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Hancock, I.; Mewes, H.W.; Mayer, K.H.
submitted to the Protein Sequence Database, February 1999
A:Reference number: 215420
A:Accession: T05646
A:Molecule type: DNA
A:Residues: 1-111 <BEV>
A:Cross-references: EMBL:AL035531
A:Experimental source: cultivar Columbia; BAC clone F20D10
C:Genetics:
A:Map position: 4
A:Introns: 139/2; 575/3
A:Note: F20D10.310

Query Match 20.6%; Score 77.5; DB 2: Length 111;
Best Local Similarity 29.0%; Pred. No. 12;
Matches 20: Conservative 13: Mismatches 25: Indels 11: Gaps 2:

QY 2 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLRS-GLKPTPSLKPD-----VPP 34
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 5 PQTGK----KVRNNSGCTQTKFAHRYSGPVVSLSDNTLSSGSLSGDYSNVTVMIFP 60
QY 55 KTSFAPLSL 63
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 61 TPNQPMAT 69

RESULT 4
T47463
serine/proline-rich protein - Arabidopsis thaliana
N:Alternate names: protein T14D3.176
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 26-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 26-Apr-2000
R:Accession: T47463
R:Author: N.; Bangert, S.; Wiedemann, R.; Voes, H.; Unseld, M.; Mewes, H.W.; Lewicki, K.
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224467
A:Accession: T47463
A:Molecule type: DNA
A:Status: preliminary
A:Cross-references: EMBL:AL138469
A:Residues: 1-175 <JCR>
A:Experimental source: cultivar Columbia; BAC clone T14D3
C:Genetics:
A:Map position: 3
A:Note: T14D3.176

Query Match 20.5%; Score 77; DB 2: Length 175;
Best Local Similarity 27.0%; Pred. No. 13;
Matches 20: Conservative 14: Mismatches 25: Indels 14: Gaps 2;

QY 1 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLRS-GLKPTPSLKPDV----HPKP 56
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 27 PAPSLDLADSLIHASPPS-----K-GSHNSPAESPIEYSFPEFETHSFSPSP 76
QY 57 SPAPLSISMKNDA 70
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I

```

QY 77 ANSPSVSPHPLNLS 90

#### RESULT 5

ER6255  
Hypothetical protein (imported) - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
R:Accession: ER6255  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al-  
Chal, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar  
ansen, N.P.; Hughes, B.; Ruizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kir  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz  
Rizzo, M.; Rooney, T.; Rowley, D.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A8614; MUID:21016719; PMID:11130712  
A:Accession: ER6255  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-744 <STO>  
A:Cross-references: GB:AE005172; NID:q157926; PIDN:AAC17609.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 20.3%; Score 76.5; DB 2: Length 744;  
Best Local Similarity 30.3%; Pred. No. 9.6;  
Matches 20: Conservative 14: Mismatches 23: Indels 9: Gaps 2;

QY 1 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLRS-GLKPTPSLKPDVPPKPSFAP 60  
 I

Db 398 PPP-----LVVSPPPPSKKMS--PTVRYSPPPPPSKKSPSVRYNSPPPPPSK 448  
 I

QY 61 LSTSMK 66  
 I

Db 445 MSFSVR 454  
 I

#### RESULT 6

T3103  
c10-1 protein - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 27-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jun-2000  
C:Accession: T3103  
C:Author: N.; Havelka, J.  
submitted to the EMBL Data Library, May 1998  
A:Description: The sequence of C. elegans cosmid C37F5.  
A:Reference number: 221283  
A:Accession: T3103  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-393 <ML>  
A:Cross-references: EMBL:AF067606; PIDN:AAC17530.1; GSPDB:GN00022; CESP:C37F5.1  
A:Experimental source: strain Bristol N2; clone C37F5  
C:Genetics:  
A:Map position: 4  
A:Introns: 94/3; 188/2; 330/1  
C:Superfamily: clk-1 transforming protein; ets DNA-binding domain homology  
F:25-105/Domain: ets DNA-binding domain homology <ETS>

Query Match 20.1%; Score 75.5; DB 2: Length 393;  
Best Local Similarity 30.9%; Pred. No. 6;  
Matches 20: Conservative 4: Mismatches 32: Indels 31: Gaps 5;

QY 1 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLRS-GLKPTPSLKPDV----VDSIQVHS-SQPS-----GQAVTVSRQPSLNAYNSL 37  
 I

Db 151 PPPPPONRGNTDFSAISLLGDSPTTHSVSTPDSVCSPPSSSVASSATPSTSSPVD 210  
 I

QY 38 TRSGLKRTPSLKPD-----VPPKPSFAPLSLSTSMKPN 68  
 I





## RESULT 15

T03454

ALR protein - human

C:Species: Homo sapiens (man)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_exchange 21-Jul-2003

C:Accession: T03454

R:Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, P.; Bruck, T.; Ballapalli, R.; Yanov

Oncogene 15, 549-560, 1997

A:Title: Structure and expression pattern of human ALR, a novel gene with strong homology

A:Reference number: Z14954; EMBL:Z14954; F010:Z14954

A:Accession: T03454

A&gt;Status: preliminary; translated from Gb/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-5262 &lt;PRA&gt;

A:Cross-references: EMBL:AF010403; NID:q2356234; PDB:1AAC51734.1; F010:q2356285

C:Genetics:

A:Gene: ALR

A:Map position: 12

C:Superfamily: human ALR protein

C:Keywords: alternative splicing

Query Match

Best local similarity 34.7% Prod. No. 20-023

Matches 25, Conservative 6, Mismatches 44, Indels 17, Gaps 4

QY

12 QVHSSQPSGAVTSVPSQSLNAINSTHSLKAT-----LSLRDYP---PKP --- 56

Db

2230 ELKAKVPSQPPNPNVRSCTGTAPVG-TTSPMRFTFOAVGFPSLKPVPQPSLPDHPGICN LDR

QY

57 -SPAP-STSMPK 67

Db

2289 SHFSPGFTLGKP 2300

Search completed: September 30, 2003, 16:35:30

Job time : 55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model:

Run on: September 30, 2003, 15:15:02 : Search time 38 Seconds  
(without alignment) : 34 hrs / sec  
69,103 Million cell updates/sec

Title: US-09-856-681-4  
Perfect score: 376  
Sequence: 1 PPAFQVDSIQVHSQSG.....PPKPSAPLSTSMKNDKAT 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127653 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 12765

Minimum DB seq length: 0

Maximum DB seq length: 260000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 4% summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length DB	ID	Description
1	376	100.0	1030	1	SMGA_HUMAN
2	87	23.1	961	1	FGD1_HUMAN
3	86.5	23.0	564	1	ZYX_MOUSE
4	86	22.9	862	1	MAK3_RAT
5	77	20.5	960	1	FGD1_MOUSE
6	75.5	20.1	397	1	GAT5_HUMAN
7	75.5	20.1	477	1	MAZ_HUMAN
8	75.5	20.1	5147	1	PCO2_HUMAN
9	74	19.7	452	1	HIS7_HYPR
10	74	19.7	894	1	MAK3_HUMAN
11	73.5	19.5	5085	1	PCO2_RAT
12	72.5	19.3	628	1	V70K_TYRV
13	72.5	19.0	446	1	TEB3_MOUSE
14	71.5	19.0	669	1	SEEL_HUMAN
15	71.5	19.0	1004	1	PRC1_HUMAN
16	71.5	19.0	1522	1	PSI1_SCHRO
17	71	18.9	428	1	ELK1_HUMAN
18	70.5	18.8	344	1	Z1PA_SHEON
19	70	18.6	715	1	PRK1_SCHRO
20	70	18.6	1012	1	PRC1_MOUSE
21	69.5	18.5	331	1	MAZ_MOUSE
22	69.5	18.5	477	1	MAZ_MOUSE
23	69.5	18.5	1125	1	MAP4_MOUSE
24	58.5	18.2	625	1	R101_YEAS
25	58.5	18.2	743	1	TEB3_HUMAN
26	68	18.1	737	1	SKN1_CANAL
27	68	18.1	812	1	NAH2_HUMAN
28	68	18.1	1259	1	AC12_HUMAN
29	67.5	18.0	429	1	ELK1_MOUSE
30	67.5	18.0	525	1	CO2A_HUMAN
31	67.5	18.0	629	1	V70K_TYMYA
32	67.5	18.0	629	1	V70K_TYMYC
33	67.5	18.0	5173	1	MUC2_HUMAN

34	67	17.8	315	1	YK04_CAREL
35	67	17.8	529	1	DNB2_ADE05
36	67	17.8	538	1	KNC0_YEAST
37	67	17.8	779	1	SRP_PROMO
38	67	17.8	813	1	NAH2_RAT
39	67	17.8	1078	1	S24A_HUMAN
40	67	17.8	1152	1	VSP4_HUMAN
41	67	17.8	2035	1	HFC1_HUMAN
42	66.5	17.7	563	1	WAP_HUMAN
43	66.5	17.7	837	1	RCO1_YEAST
44	66	17.6	295	1	PRB_MYCLE
45	66	17.6	433	1	HXB3_MOUSE

ALIGNMENTS

RESULT 1	
SMGA_HUMAN	
AC Q9H2E6: Q9H2E6; STANDARD: PRT: 1030 AA.	
DI 15-SEP-2003 (rel. 42, Created)	
DI 15-SEP-2003 (rel. 42, Last sequence update)	
DI 15-SEP-2003 (rel. 42, Last annotation update)	
DI Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1) (SPW6A-1)	
DI SEMATA OR KIAA1143	
DI Homo sapiens (Human)	
DI Eukaryota; Metazoa; Chordata; Vertebrata; Puteleostomi;	
DI Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
DI NCBI TaxID: 9606;	
DI 1	
DI SEQUENCE FROM N.A., AND INTERACTION WITH EVL.	
DI MEDLINE-20564339; PubMed 10993894;	
DI Klosterman A., Lutz H., Sertler F., Beil C.	
DI "the orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal tyrosine-like domain."	
DI J. Biol. Chem. 275:35647-35653(2000);	
DI 2.	
DI SEQUENCE FROM N.A.	
DI TISSUE=Brain;	
DI MEDLINE-20181126; PubMed-10718198;	
DI Nagase T., Kikuo R., Ishikawa K.-I., Hirosewa M., Ohara O.;	
DI "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro."	
DI DNA Res. 7:65-73(2000);	
DI 1- FUNCTION: Can act as repulsive axon guidance cues. May play a role in channeling sympathetic axons into the sympathetic chains and controlling the temporal sequence of sympathetic target innervation (By similarity).	
DI 1- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.	
DI 1- SUBCELLULAR LOCATION: Type I membrane protein.	
DI 1- ALTERNATIVE PRODUCTS:	
DI Event=Alternative splicing; Named isoforms=2;	
DI Name=1;	
DI IsoId-Q9H2E6-1; Sequence=Displayed;	
DI Name=2;	
DI IsoId-Q9H2E6-2; Sequence=VSP_007113;	
DI Note=No experimental confirmation available;	
DI 1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.	
DI 1- SIMILARITY: Contains 1 Sema domain.	
DI This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by not-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	

```

DR EMBL: AF279656: AAC29378.1: 1:
DR EMBL: AB037789: BAA92506.1: ALI_INI1.
DR GenBank: HGNC:10738: SEMA6A.
DR MIM: 605885.
DR InterPro: IPR003659: Flexin-like.
DR InterPro: IPR00627: SEMA.
DR Pfam: PF0403: SEMA; 1.
DR SMART: SM00423: PSI; 1.
DR SMART: SM00630: SEMA; 1.
DR Signal: Transmembrane; Multimeric family; Neurogenesis; Glycoprotein;
KW Developmental protein; Alternative splicing.
FT SIGNAL 1 18
FT CHAIN 19 1030
FT DOMAIN 19 649
FT TRANSMEM 650 670
FT DOMAIN 671 1030
FT DOMAIN 56 491
FT DOMAIN 792 819
FT CARBOHYD 33 33
FT CARBOHYD 49 49
FT CARBOHYD 55 55
FT CARBOHYD 282 282
FT CARBOHYD 434 434
FT CARBOHYD 462 461
FT VARSPIC 575 576
FT
FT
SQ SEQUENCE 1030 AA: 114366 MW: A57879C10AEC4534 QRC64:
Query Match 100.0% Score 376; DB 1: Length 1030;
Best Local Similarity 100.0%; Prod. No. 1.6e-28;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY : PPPAPQVDSIQVHSSQPSQVTSQPSLNAYNSLRSLKRTSLKPDVDPKPSFAP 60
DB 359 PPPAPQVDSIQVHSSQPSQVTSQPSLNAYNSLRSLKRTSLKPDVDPKPSFAP 1012
OY 61 LSTSMKPNDACT 72
DB 1019 LSTSMKPNDACT 1030
RESULT 2
FCGL_HUMAN
ID FCGL_HUMAN STANDARD: PRT: 961 AA.
AC P98174: Q8N4D9;
DI 01-OCT-1996 (Rel. 34, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Putative Rho/Rac guanine nucleotide exchange factor (Rho/Rac GEF)
DE (Faciogenital dysplasia protein).
GN FCGL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=45042764; PubMed=7951831;
RA Pastoris N.G., Cadle A., Logie L., Porteous M.E.M., Schwartz C.E.,
RA Stevenson R.E., Glover T.W., Wilroy R.S., Gorski J.,
RA "Isolation and characterization of the faciogenital dysplasia
RA (Aarskog-Scott syndrome) gene: a putative Rho/Rac guanine nucleotide
RA exchange factor."
RL Cell 79:669-678(1994).
RN 2;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=22388257; PubMed=72477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins L., Wagner L., Shenmen C.M., Scheller G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Fatmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Hoxaldo M.P., Casavant I.L., Scheetz T.E.,
RA Bronstein M.J., Gsdi T.H., Toshiyuki S., Carninci P., Prange C.,
RA Kaha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Sosak S.A., McEwan F.J., McKernan K.J., Malek J.A., Gnanaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Khitina M., Madan A., Young A.C., Shevchenko Y., Houliard G.G.,
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood C., Schmitz J., Myers R.M.,
RA Butterfield V.S.N., Krzywicki M.I., Skelski G., Smellus D.E.,
RA Schmeich A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 3;
RP VARIANT AAS HIS-522.
PX MEDLINE=20546218; PubMed=11093277;
RA Schwartz C.E., Gillespie-Kaeschach G., May M., Cappa M., Gorski J.L.,
RA Steindl K., Neri G.;
RT "Two novel mutations confirm FCGL is responsible for the Aarskog
RT syndrome."
RL Eur. J. Hum. Genet. 8:859-874(2000).
RN 4;
RP VARIANT AAS GUN-610.
RX MEDLINE=20389563; PubMed=10530571;
RA Orlicco A., Galli L., Falciani M., Bracci M., Cavaliere M.L.,
RA Kinaldi M.M., Musacchio A., Sorrentino V.;
RT "A mutation in the pleckstrin homology (PH) domain of the FCGL gene in
RT an Italian family with faciogenital dysplasia (Aarskog-Scott
RT syndrome)."
RL FEBS Lett. 478:215-220(2000).
RN 5;
RP FUNCTION: ACTIVATES THE RAS-LIKE FAMILY OF RHO- AND RAC PROTEINS
RP BY EXCHANGING BOUND GDP FOR FREE GTP.
CC 1- SOURCE: LOCATED IN THE CYTOSOL (By similarity).
CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC 1- TISSUE SPECIFICITY: Expressed in fetal heart, brain, lung, kidney
CC and placenta, less expressed in liver; adult heart, brain, lung,
CC pancreas and skeletal muscle.
CC 1- DISEASE: Defects in FCGL are the cause of Aarskog-Scott syndrome
CC (AAS) [MIM:305400]. This faciogenital dysplasia is a rare,
CC multisystemic disorder characterized by disproportionately short
CC stature, and by facial, skeletal, and urogenital anomalies.
CC 1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC 1- SIMILARITY: Contains 2 PH domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
RN 6;
RP FMBL: J11690; AAS57064.1;
RL FMBL: RC34530; RAH34530.1;
CC HSSP: Q07889; IAME.
DR GenBank: HGNC:3663; FCGL.
DR MIM: 305400; 1.
DR GO: GO:0005085; E-cadherin nucleotide exchange factor activity; IAS.
DR GO: GO:0007275; P:development; TAS.
DR GO: GO:0007397; P:histogenesis and organogenesis; TAS.
DR GO: GO:0007165; P:sigal transduction; TAS.
DR InterPro: IPR001331; GDS_CDC24.
DR InterPro: IPR001349; PH.
DR InterPro: IPR000219; RhoGEF.
DR InterPro: IPR000306; Znf_FYVE.
DR Pfam: PF01363; FYVE; 1.
DR Pfam: PF00169; PH; 2.
DR Pfam: PF00621; RhoGEF; 1.
DR SMART: SM00064; FYVE; 1.
DR SMART: SM00233; PH; 2.

```





CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF12224; AAK532.4.1; -  
 CC HSSP: P24941; 180H  
 CC InterPro: IPR000180; Glnase.  
 CC InterPro: IPR000719; Pro\_kinase.  
 CC InterPro: IPR002290; Ser\_thr\_kinase.  
 CC Pfam: PF00780; CNH; 1.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD003001; Prot\_kinase; 1.  
 CC SMART: SM00036; CNH; 1.  
 CC SMART: SM00320; STK; 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE: PS00108; PROTEIN\_KINASE\_ST; FALSE\_NEG.  
 CC ATP-binding: Transferrase; Serine/threonine-protein kinase.  
 CC NON\_TER 1 1  
 CC DOMAIN 5 262 PROTEIN KINASE.  
 CC FT DOMAIN 530 842 CNH.  
 CC FT NP\_BIND 11 19 ATP (BY SIMILARITY).  
 CC FT BINDING 34 34 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 125 125 BY SIMILARITY.  
 CC SEQUENCE 862 AA: 97390 MW: 58013AC3B0A3287F CRC64;  
 Query Match 22.9% Score 46; DB 1; Length 862;  
 Best Local Similarity 34.7%; Pred. No. 0.72;  
 Matches 25; Conservative 5; Mismatches 27; Indels 14; Gaps 3;  
 QY : PPPAPQVDS:TCV-----HSSQFSGQAVTSKQSLNAYNSLTSGLKRTFSIKHWPEK 56  
 DB 400 PPLPPPKXSTSLPQDTHSSDSNQS:TTAKCS-----SSPAKSHVPPPPPP 449  
 QY 57 STAPLSTSMKPN 58  
 DB 450 RLPPQKPAVIGN 46;  
 RESULT 5  
 FGBL\_MOUSE  
 ID FGBL\_MOUSE STANDARD; PRT; 960 AA.  
 AC P52734;  
 DT 01-OCT-1996 (Rel. 34; Created)  
 DT 01-OCT-1996 (Rel. 34; Last sequence update)  
 FT 26-FEB-2003 (Rel. 41; Last annotation update)  
 FE putative rho/rac guanine nucleotide exchange factor (GEF)/rac GEF)  
 DE (Raclophenital dysplasia protein homolog).  
 GN FGBL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96081343; PubMed=8538076;  
 RX Pasteris N.G., de Gouyon B., Cadie A.R., Campbell K., Horan G.E.,  
 RA Gorski J.;  
 RT Cloning and regional localization of the mouse laclophenital  
 RI dysplasia (Fgbl) gene.;  
 RL Mamm. Genome 6:558-561(1995).  
 CC -!- FUNCTION: ACTIVATES THE RAS-LIKE FAMILY OF RHO AND RAC PROTEINS  
 CC BY EXCHANGING BOUND GDP FOR FREE GTP.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Contains 1 Dbl-homology (DB) domain.  
 CC -!- SIMILARITY: Contains 2 FVE-Gesalts.  
 CC -!- SIMILARITY: Contains 1 FVE-type zinc finger.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U22325; AAA96001.1; -  
 CC HSSP: Q07889; LAME.  
 CC MCE: MCE104566; Fgd1.  
 CC InterPro: IPR001331; GDS\_GDC24.  
 CC InterPro: IPR001849; PH.  
 CC InterPro: IPR000219; RhogEF.  
 CC InterPro: IPR000306; Znf\_FVVE.  
 CC Pfam: PF01363; FVVE; 1.  
 CC Pfam: PF01169; PH; 2.  
 CC Pfam: PF06621; RhogEF; 1.  
 CC SMART: SM00064; FVVE; 1.  
 CC SMART: SM00233; PH; 2.  
 CC SMART: SM00325; RhogEF; 1.  
 CC PROSITE: PS00741; DH; 1; FALSE\_NEG.  
 CC PROSITE: PS00010; DH; 2; 1.  
 CC PROSITE: PS00063; PH\_DOMAIN; 2.  
 CC PROSITE: PS0178; ZNFVVE; 1.  
 CC KW Guanine-nucleotide releasing factor; zinc-finger; Repeat.  
 CC FT DOMAIN 372 560 DH.  
 CC FT DOMAIN 7 330 PRO-RICH.  
 CC FT SLIT 171 179 SH3-BINDING (POTENTIAL).  
 CC FT SLIT 179 187 SH3-BINDING (POTENTIAL).  
 CC FT DOMAIN 589 688 PH 1.  
 CC FT ZNFING 729 789 FIVE-TYPE.  
 CC FT DOMAIN 820 920 PH 2.  
 CC SEQUENCE 960 AA: 106477 MW: 4C1B84DF450FC51 CRC64;  
 Query Match 20.5% Score 77; DB 1; Length 960;  
 Best Local Similarity 31.8%; Pred. No. 5.9;  
 Matches 21; Conservative 5; Mismatches 24; Indels 16; Gaps 2;  
 QY 2 PPAPQVDS:TCV-----HSSQFSGQAVTSKQSLNAYNSLTSGLKRTFSIKHWPEK 63  
 DB 127 PEGPQF-----KSDPGFTEIPGRP-----SPLKRAPGPKPQVPPKPSYLM 170  
 QY 62 SLSMKP 67  
 DB 271 PVIPT 176  
 RESULT 6  
 GAT5\_HUMAN  
 ID GAT5\_HUMAN STANDARD; PRT; 397 AA.  
 AC Q9HXX5;  
 DT 28-FEB-2003 (Rel. 41; Created)  
 DT 26-FEB-2003 (Rel. 41; Last sequence update)  
 DT 28-FEB-2003 (Rel. 41; Last annotation update)  
 DE Transcription factor GATA-5 (GATA binding factor-5).  
 GN GATA5  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21638749; PubMed=11780052;  
 RA DeCunha P., Matthews D.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagquell C.L.,  
 RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley C.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Cledg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Gratham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J., Heath P.D., Ho S., Holden J.L., Howden P.J.,

```

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Ray M.P., Kimberley A.M., King A., Knaighs A., Laird G.K., Lawler S.,
RA Lehaestath M.H., Loversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson J.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.L.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showknee R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund G., Steward C.A., Stilson J.E.,
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Treman A.C., Vaudin R., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D., Williams E., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Dartin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RA "The DNA sequence and comparative analysis of human chromosomes 20-21
RL Nature 414:65-87;(2001).
CC -1- FUNCTION: BINDS TO THE FUNCTIONALLY IMPORTANT DEF-1 NUCLEAR
CC PROTEIN BINDING SITE IN THE CARDIAC-SPECIFIC SODIUM/CARDIAC TROPONIN
CC C TRANSCRIPTIONAL ENHANCER. MAY PLAY AN IMPORTANT ROLE IN THE
CC TRANSCRIPTIONAL PROGRAM(S) THAT UNDERLIES SMOOTH MUSCLE CELL
CC DIVERSITY (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 2 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: A249627; GAC36001.1; .
CC HSSP: P17679; IGNF.
CC Genew: HGNC:15802; GATA5.
CC InterPro: IPR006679; Znf_GATA.
CC Pfam: PF05349; GATA-N; .
CC Pfam: PF00320; GATA; 2.
CC PRINTS: PR00619; GATAZNFINGER.
CC SMART: SM00401; ZF-GATA; 2.
CC PROSITE: PS00344; GATA_ZNFINGER_1; 2.
CC PROSITE: PS00114; GATA_ZNFINGER_2; 2.
CC K1 Transcription factor; Activator; DNA-binding; Zinc-finger;
CC K2 Nuclear protein.
CC FT 2N.FING 189 213 GATA-TYPE.
CC FT 2N.FING 243 267 GATA-TYPE.
CC SQ SEQUENCE 397 AA: 41299 MW: 50580208695057 CRK64:
Query Match 20.1% Score 75.5; DB 1; Length 397;
Best Local Similarity 29.1%; Pred. No. 3;
Matches 22; Conservative 8; Mismatches 36; Indels 5; Gaps 1;
QY 1 PPAPORVLISQVHSSQP-----SQAVTVSPQPSLNAYNSLRGLKRRLYSKPL 51
DB 277 PRPLAMKKEISQIRKKPKPIAKRGSSGSTRNASASPSAVASTOSAKESIASPE 224
QY 52 VPKPSFAPLSTSMK 66
DB 337 VQPSRSMARQASQCE 351

```

## RESULT 7

```

MAZ_HUMAN
ID MAZ_HUMAN STANDARD; PRT: 477 AA.
AC P56270; Q15763; Q99443;
DT 15-JUL-1993 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE MYC-associated zinc finger protein (MAZI) (Purine binding
DE Transcription factor) (Pur-1) (ZF87) (ZF87).
GN MAZ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; CranLata; Vertebrata; Euteleostomi;

```

```

CC Mammalia; Butteria; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=92366475; PubMed=3502157;
CC Bressane S.A., Asselin C., Patel A.J., Marcu K.B.;
CC "MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences
CC regulating transcriptional initiation and termination.";
CC Proc. Natl. Acad. Sci. U.S.A. 89:7452-7456(1992).
CC [2]
CC SEQUENCE FROM N.A.
CC TISSUE=Carcinoma;
CC MEDLINE=92323279; PubMed=1567856;
CC Fyfe G.J., Mether K.H., Hall D.J.;
CC "Isolation of a novel cDNA encoding a zinc-finger protein that binds
CC to two sites within the c-myc promoter.";
CC Biochemistry 31:4102-4110(1992).
CC [3]
CC SEQUENCE FROM N.A.
CC TISSUE=Pancreatic islets;
CC MEDLINE=9642551; PubMed=8431693;
CC Tsutsui H., Sakatsune O., Itakura K., Yokoyama K.K.;
CC "Members of the MAZ family: a novel cDNA clone for MAZ from human
CC pancreatic islet cells.";
CC Biochem. Biophys. Res. Commun. 226:801-809(1995).
CC [4]
CC SEQUENCE FROM N.A.
CC MEDLINE=96224025; PubMed=8626793;
CC Parks C.L., Shonk T.;
CC "The serotonin 1a receptor gene contains a TATA-less promoter that
CC responds to MAZ and Sp1.";
CC J. Biol. Chem. 271:4417-4430(1996).
CC [5]
CC SEQUENCE FROM N.A.
CC TISSUE=Myophoblastoma;
CC MEDLINE=98352105; PubMed=9685418;
CC Song J., Karakami H., Tsutsui H., Iano X., Matsumura M., Itakura K.,
CC Kanazawa I., Sun K., Yokoyama K.K.;
CC "Genomic organization and expression of a human gene for Myc-
CC associated zinc finger protein (MAZ).";
CC J. Biol. Chem. 273:20603-20614(1998).
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR WITH DUAL ROLES
CC IN TRANSCRIPTION INITIATION AND TERMINATION. BINDS TO TWO SITES,
CC MEL1 AND MEL2, WITHIN THE C-MYC PROMOTER HAVING GREATER
CC AFFINITY FOR THE FORMER. ALSO BINDS TO MULTIPLE G/C-RICH SITES
CC WITHIN THE PROMOTER OF THE SPI FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL
CC MUSCLE, AND PANCREAS. SEEMS NOT TO BE EXPRESSED IN KIDNEY.
CC -1- SIMILARITY: Contains 6 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M4046; .; NOT_ANNOTATED_CDS.
CC EMBL: M93339; .; NOT_ANNOTATED_CDS.
CC EMBL: D8513; BAA12728.1; ALT_INIT.
CC EMBL: U3619; AAB0412.1; ALT_INIT.
CC EMBL: AB017335; BAA33064.1; .
CC FIC: A42170; A42170.
CC TRANSFAC: T06490; .
CC TRANSFAC: T02105; .
CC Genew: HGNC:6914; MAZ.
CC NM_200999; .
CC GO:0005367; P:transcription initiation from Pol II promoter; TAS.
CC GO:0006369; P:transcription termination from Pol II promoter; TAS.
CC InterPro: IPR007087; Znf_C2H2.
CC Pfam: PF00096; Zf-C2H2; 5.

```

```

DR PRODOM: P0000003: Znf_C2H2: 1.
DR SMART: SM00355: ZNF_C2H2: 6
DR PROSITE: PS00628: ZINC_FINGER_C2H2_1: 5.
DR PROSITE: PS00157: ZINC_FINGER_C2H2_2: 5.
KW Transcription regulation; Zinc-finger; Metal-binding; Metal-binding; DNA-binding;
KW RNA-binding; Repeat; Nuclear protein.
FT ZNF-FING 190 212 C2H2-TYPE 1.
FT ZNF-FING 279 301 C2H2-TYPE 2.
FT ZNF-FING 307 329 C2H2-TYPE 3.
FT ZNF-FING 337 360 C2H2-TYPE 4.
FT ZNF-FING 366 388 C2H2-TYPE 5.
FT ZNF-FING 392 413 C2H2-TYPE 6.
FT DOMAIN 96 108 POLY-ALA.
FT DOMAIN 133 139 POLY-PRO.
FT DOMAIN 157 161 POLY-ALA.
FT DOMAIN 245 249 POLY-GLY.
FT DOMAIN 435 445 POLY-ALA.
FT DOMAIN 455 459 MISSING (IN REF. 3).
FT CONFLICT 431 431 L -> M (IN REF. 2 AND 4).
FT CONFLICT 443 447 MISSING (IN REF. 3).
SQ SEQUENCE 477 AA: 48607 MW: 60460720.46825 CHC64;

Query Match 20.1%; Score 75.5; DS 1; Length 477;
Best Local Similarity 25.8%; Proc. No. 3.7;
Matches 23; Conservative 15; Mismatches 32; Indels 19; Gaps 3;

QY 1 PPPAPQ-----RVDSICV-----HSSQPSGCAVTSPQSLNAYNS-TRSGDK 43
DB 59 PPPTPOAARPELOVDLLEVLAAQESRAAANAAMAAVAPPAAPASVETALK 128

QY 44 RTFSLKPVPKPKSPAPLETSMKFNDACT 72
DB 129 QPPA--PPPPPPVSPAPAAAPASAAAT 155

RESULT 8
PCLO_HUMAN STANDARD: PRT: 5147 AA.
AC Q9Y6V0; Q43379; Q60305; Q9Y6V0; Q9Y6V2; Q9Y6V3;
DT 28-FEB-2003 (Ref. 41, Created)
DI 29-FEB-2003 (Ref. 41, Last sequence update)
DI 15-SEP-2003 (Ref. 42, Last annotation update);
DE Piccolo protein (Acronin) (Fragments).
GN PCLO OR ACZ OR KIAA0554.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-759 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94439764; PubMed=10509962;
RA Wang X., Kibschall M., Labe M.M., Lippert R., Petrasch-Pareus E.,
RA Kiliama M.W.;
RA "Acronin, a 550-kD putative scaffolding protein of presynaptic active
RA zones, shares homology regions with rim and bassoon and binds
RA profilin";
RC J. Cell Biol. 147:151-162(1999).
RN [2]
RP SEQUENCE OF 552-4404 FROM N.A.
RX Kraemer J., Wollam C., Woldmann P., McGrath B.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Takaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1999).
RN [4]

```

```

RP SEQUENCE OF 4405-4419 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22589257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derce J.G.,
RA Klapper R.D., Collins R.S., Wagner L., Shenmen C.M., Schuler G.E.,
RA Altshul S.F., Zeng H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins K.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong J.,
RA Stapleton M., Soares M.B., Boraide M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.S., Udell T.R., Toshiyuki S., Carinci P., Prasad C.,
RA Raha S.S., Loquiano N.A., Peters G.J., Abramson R.D., Mullahy S.G.,
RA Sczak S.A., McMan P.J., McKernan K.J., Malek J.A., Gharatne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay I.J., Halyk S.K.,
RA Villan D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahevi J., Hellon E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Brakesley A.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Hutterfield Y.S.N., Kizyewski M.I., Skalska J., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marita M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 4405-5147 FROM N.A.
RC Kallunki J., Elliott G.;
RX Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May act as a scaffolding protein involved in the
CC organization of synaptic active zones and in synaptic vesicle
CC trafficking (by similarity).
CC -!- SUBUNIT: Interacts with Rabac1/Pral and profilin (by similarity).
CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC synaptic junctions (by similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing. Named isoforms-2;
CC Comment-Additional isoforms seem to exist;
CC Name-1;
CC IsoId-Q9Y6V0-1; Sequence-Displayed;
CC Name-2;
CC IsoId-Q9Y6V0-2; Sequence-VSP_003923, VSP_003924, VSP_003925,
CC VSP_003926, VSP_003927;
CC Note-NCBI experimental confirmation available;
CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC phospholipids. Calcium binds with low affinity but with high
CC specificity and induces a large conformational change.
CC -!- SIMILARITY: Contains 2 C2 domains.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to: license@sib-sb.ch).
CC -----
CC ENBL: Y19188; CAB60727.1;
CC DR ENBL: A004903; AAD20936.1;
CC DR ENBL: A004886; AAD21784.1;
CC DR ENBL: AB011131; BAA25485.1;
CC DR ENBL: BC001304; AAB01304.1;
CC DR ENBL: AC004082; AAB97937.1;
CC DR PIR: T06634; T06634;
CC HSPSP: P04410; IAZ5;
CC Genew: HGNC:13406; PCLO;
CC MIM: 604918;
CC GO: GO:0005856; Cytoskeleton; NAS;
CC GO: GO:0045202; Cytoskeletal junction; ISS;
CC GO: GO:0005509; F-actin ion binding activity; ISS;
CC GO: GO:0005544; F-actin-dependent phospholipid binding activity; ISS;
CC GO: GO:0005522; Profilin binding activity; ISS;
CC GO: GO:0007010; P-cytoskeleton organization and biogenesis; ISS;
CC GO: GO:0016080; P-synaptic vesicle targeting; ISS;

```



phosphoprotein.  
 CC -1- CoPACTOR: Magnesium.  
 CC -1- SUBUNIT: Interacts with SHG2. Interaction appears to require  
 CC MAPK3-mediated JNK activation (By similarity).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing. Named Isoforms:  
 CC Name=1:  
 CC IsoId=Q8IVH8-1; Sequence=VSP\_007054;  
 CC Note=No experimental confirmation available;  
 CC Name=2;  
 CC IsoId=Q8IVH8-2; Sequence=VSP\_007054;  
 CC Name=3;  
 CC IsoId=Q8IVH8-3; Sequence=VSP\_007054;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed in all tissue  
 CC examined, with high levels in heart, brain, placenta, skeletal  
 CC muscle, kidney and pancreas and lower levels in lung and  
 CC liver.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC STR20 SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 CNH domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; AF000145; AAC15472.1; .  
 DR EMBL; AF445413; AAN75849.1; JOINED.  
 DR EMBL; AF445385; AAN75849.1; JOINED.  
 DR EMBL; AF445386; AAN75849.1; JOINED.  
 DR EMBL; AF445387; AAN75849.1; JOINED.  
 DR EMBL; AF445388; AAN75849.1; JOINED.  
 DR EMBL; AF445389; AAN75849.1; JOINED.  
 DR EMBL; AF445390; AAN75849.1; JOINED.  
 DR EMBL; AF445391; AAN75849.1; JOINED.  
 DR EMBL; AF445392; AAN75849.1; JOINED.  
 DR EMBL; AF445393; AAN75849.1; JOINED.  
 DR EMBL; AF445394; AAN75849.1; JOINED.  
 DR EMBL; AF445395; AAN75849.1; JOINED.  
 DR EMBL; AF445396; AAN75849.1; JOINED.  
 DR EMBL; AF445397; AAN75849.1; JOINED.  
 DR EMBL; AF445398; AAN75849.1; JOINED.  
 DR EMBL; AF445399; AAN75849.1; JOINED.  
 DR EMBL; AF445400; AAN75849.1; JOINED.  
 DR EMBL; AF445401; AAN75849.1; JOINED.  
 DR EMBL; AF445402; AAN75849.1; JOINED.  
 DR EMBL; AF445403; AAN75849.1; JOINED.  
 DR EMBL; AF445404; AAN75849.1; JOINED.  
 DR EMBL; AF445405; AAN75849.1; JOINED.  
 DR EMBL; AF445406; AAN75849.1; JOINED.  
 DR EMBL; AF445407; AAN75849.1; JOINED.  
 DR EMBL; AF445408; AAN75849.1; JOINED.  
 DR EMBL; AF445409; AAN75849.1; JOINED.  
 DR EMBL; AF445410; AAN75849.1; JOINED.  
 DR EMBL; AF445411; AAN75849.1; JOINED.  
 DR EMBL; AF445412; AAN75849.1; JOINED.  
 DR EMBL; AF445413; AAN75849.1; JOINED.  
 DR EMBL; AF445385; AAN75850.1; JOINED.  
 DR EMBL; AF445386; AAN75850.1; JOINED.  
 DR EMBL; AF445387; AAN75850.1; JOINED.  
 DR EMBL; AF445388; AAN75850.1; JOINED.  
 DR EMBL; AF445390; AAN75850.1; JOINED.  
 DR EMBL; AF445391; AAN75850.1; JOINED.  
 DR EMBL; AF445392; AAN75850.1; JOINED.  
 DR EMBL; AF445393; AAN75850.1; JOINED.  
 DR EMBL; AF445394; AAN75850.1; JOINED.  
 DR EMBL; AF445395; AAN75850.1; JOINED.  
 DR EMBL; AF445397; AAN75850.1; JOINED.  
 DR EMBL; AF445398; AAN75850.1; JOINED.  
 DR EMBL; AF445399; AAN75850.1; JOINED.

DR EMBL; AF445400; AAN75850.1; JOINED.  
 DR EMBL; AF445401; AAN75850.1; JOINED.  
 DR EMBL; AF445402; AAN75850.1; JOINED.  
 DR EMBL; AF445403; AAN75850.1; JOINED.  
 DR EMBL; AF445404; AAN75850.1; JOINED.  
 DR EMBL; AF445405; AAN75850.1; JOINED.  
 DR EMBL; AF445406; AAN75850.1; JOINED.  
 DR EMBL; AF445407; AAN75850.1; JOINED.  
 DR EMBL; AF445408; AAN75850.1; JOINED.  
 DR EMBL; AF445409; AAN75850.1; JOINED.  
 DR EMBL; AF445410; AAN75850.1; JOINED.  
 DR EMBL; AF445411; AAN75850.1; JOINED.  
 DR EMBL; AF445412; AAN75850.1; JOINED.  
 DR EMBL; A0007684; AAF19249.1; .  
 DR HSSP; P24941; 1B36.  
 DR Genew; HGNC:6955; MAP4K3.  
 DR MIM; 694921; .  
 DR InterPro; IPR001180; Citron.  
 DR InterPro; IPR000719; Prot\_Kinase.  
 DR InterPro; IPR002230; Ser\_Thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00780; CNH; 1.  
 DR Pfam; PF00669; pkinase; 1.  
 DR ProDom; PD000001; Prot\_Kinase; 1.  
 DR SMART; SM00036; CNH; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_AIP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_SI; FALSE\_NEG.  
 DR ATP-binding; transferase; Serine/threonine-protein kinase;  
 KW Alternative splicing.  
 FT DOMAIN 16 273 PROTEIN KINASE.  
 FT DOMAIN 562 874 CNH.  
 FT NP\_3'ND 22 30 ATP (BY SIMILARITY).  
 FT BINDING 43 48 ATP.  
 FT ACT\_SITE 136 136 BY SIMILARITY.  
 FT VARSPPLIC 1 12 MNPGLSLRRNP -> MA (in isoform 2).  
 FT VARSPPLIC 352 372 /FTID-VSP\_007052.  
 FT MUTAGEN 48 48 K->E; LOSS OF KINASE ACTIVITY AND ABILITY  
 FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).  
 FT SEQUENCE 894 AA; 101315 MW; 6EH75BB34ESB73 CRC64;  
 Query Match 19.7%; Score 74; DB 1; Length 894;  
 Best Local Similarity 32.8%; Pred. No. 10;  
 Matches 2; Conservative 6; Mismatches 23; Indels 14; Gaps 3;  
 1 PFFAFQKVDST-----GVHSSQFSGQAVTVSROPSLNAYNSLZRSGLKRTPLSKPDVPPKP 56  
 432 PPLPLPKPKSLFIPQMHSHSEFNQG--TKRCP-----MSGSPAKPSQVPPRPHP 491  
 57 STAP 56  
 462 RIPP 485  
 RESULT 11  
 PCOL\_RAT  
 ID PCOL\_RAT STANDARD; PRT; 5085 AA.  
 AC Q9JKS5; Q9JL11.  
 DT 28-FEB-2003 (Rel. 41; Created)  
 DT 28-FEB-2003 (Rel. 41; Last sequence update)  
 DT 28-FEB-2003 (Rel. 41; Last annotation update)  
 DE Piccolo protein (Multidomain presynaptic cytomatrix protein).  
 GN PCLO.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 CX NCBI\_taxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH KABA1.

RX MEDLINE 20170257; PubMed-10767984;  
 RA Fenster S.D., Chung W.J., Zhai R., Cases-Landhoff C., Voss B.,  
 RA Garner A.M., Kacmpf C., Kinder S., Gundelfinger E.D., Garner C.C.,  
 RT "Piccolo, a presynaptic zinc finger protein structurally related to  
 RT bassoon",  
 RL Neuron 25:203-214 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Fenster S.D., Cases-Landhoff C., Gundelfinger E.D., Garner C.C.,  
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP CALCIUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668; ASP-4674;  
 RP VAL-4688; MET-4689; VAL-4692; SER-4693; GLN-4694; ASN-4695 AND  
 RP ALA-4694.  
 RX MEDLINE-21181819; PubMed-11285225;  
 RA Gerber S.H., Garcia J., Rizo J., Suchofsky J.D.,  
 RT "An unusual C(2)-domain in the active-zone protein piccolo:  
 RT implications for Ca(2+) regulation of neurotransmitter release",  
 RL EMBO J. 20:1605-1619 (2001).  
 CC -1- FUNCTION: May act as a scaffolding protein involved in the  
 CC organization of synaptic active zones and in synaptic vesicle  
 CC trafficking (By similarity).  
 CC -1- SUBUNIT: Interacts with Rabact/Pral and profilin.  
 CC -1- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic  
 CC junctions.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-A: Alternative splicing; Named isoforms 1-2;  
 CC Name-1:  
 CC IsoId-09JXK56-1; Sequence-Displayed;  
 CC Name-2:  
 CC IsoId-09JXK56-2; Sequence-VSP-004936; VSP-003931;  
 CC DOMAIN: C2 domain 1 is involved in binding calcium and  
 CC phospholipids. Calcium binds with low affinity but with high  
 CC specificity and induces a large conformational change.  
 CC -1- SIMILARITY: Contains 2 C2 domains.  
 CC -1- SIMILARITY: Contains 1 PDZ/CHR domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF138789; AAF07822.2; -  
 DR HSSP: P04410; IA25.  
 DR GO: GO:0045202; C:synaptic junction; IDA.  
 DR GO: GO:0005509; F:calcium ion binding activity; IDA.  
 DR GO: GO:0005544; F:calcium-dependent phospholipid binding activity; IDA.  
 DR GO: GO:0005522; F:profilin binding activity; ISS.  
 DR GO: GO:0030700; P:cytoskeleton organization and biogenesis; ISS.  
 DR GO: GO:0036080; P:synaptic vesicle targeting; NAS.  
 DR InterPro: IPR003008; C2.  
 DR InterPro: IPR001478; PDZ.  
 DR Pfam: PF00168; C2; 2.  
 DR Pfam: PF00595; PDZ; 1.  
 DR SMART: SM00239; C2; 2.  
 DR SMART: SM00228; PDZ; 1.  
 DR PROSITE: PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE: PS00004; C2\_DOMAIN\_2; 2.  
 DR PROSITE: PS0106; PDZ; 1.  
 KW Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;  
 FT Repeat, Alternative splicing; 12 x 10 AA TANDEN APPROXIMATE REPEATS OF  
 FT DOMAIN 372 491  
 FT P-A-K-P-Q-P-Q-P-X.  
 FT C4-TYPE (POTENTIAL).  
 FT C4-TYPE (POTENTIAL).  
 FT PDZ.  
 FT PDZ.  
 FT C2 DOMAIN 1.  
 FT C2 DOMAIN 2.

FT VARSPLIC 4876 4880  
 FT VARSPLIC 4881 5085  
 FT MUTAGEN 4668 4668  
 FT MUTAGEN 4674 4674  
 FT MUTAGEN 4668 4668  
 FT MUTAGEN 4688 4688  
 FT MUTAGEN 4688 4688  
 FT MUTAGEN 4689 4689  
 FT MUTAGEN 4690 4691  
 FT MUTAGEN 4692 4693  
 FT MUTAGEN 4694 4694  
 FT SEQUENCE 5085 AA; 552702 MW; 5A1BR543201A7450 CRC64;  
 Query Match 19.5%; Score 73.5; DB 1; Length 5085;  
 Best Local Similarity 26.4%; Pred. No. 81;  
 Matches 19; Conservative 8; Mismatches 20; Indels 25; Gaps 3;  
 QY 1 PPTAFQVLSIGVHSSQSSQAVTVSKOPS-NAYNSIGRSLKFTFSLKIUVPKPSFAP 60  
 ID 1 PPTAFQVLSIGVHSSQSSQAVTVSKOPS-NAYNSIGRSLKFTFSLKIUVPKPSFAP 60  
 QY 51 LS-TSMKPNQAC 71  
 ID 51 LS-TSMKPNQAC 71  
 QY 2488 IIVTQKPTQTS 2479  
 ID 2488 IIVTQKPTQTS 2479  
 RESULT 12  
 VTOK\_TYVV STANDARD; PRT: 628 AA.  
 AC P10357;  
 DT 01-MAR-1989 (rel. 10; Created)  
 DT 01-AUG-1992 (rel. 23; Last sequence update)  
 DT 01-AUG-1992 (rel. 23; Last annotation update)  
 DE 69 kDa protein.  
 CS Turnip yellow mosaic virus.  
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.  
 CC NCBI Taxid:12554;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-88289359; PubMed-3399388;  
 RA Morch M.D., Boyer J.C., Hachmi A.,  
 RT "Overlapping open reading frames revealed by complete nucleotide  
 RT sequencing of turnip yellow mosaic virus genomic RNA.",  
 RL Nucleic Acids Res. 16:6157-6173 (1988).  
 CC -1- FUNCTION: NOT KNOWN.  
 CC -1- SIMILARITY: TO 65 TO 70 KDA PROTEIN FROM OTHER TYMOVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X07441; CAA30321.1; ALT\_SEQ.  
 DR PIR: S01955; S01955.  
 DR InterPro: IPR004935; Tymo\_45\_70kDa.  
 DR Pfam: PF03251; Tymo\_45kd\_70kd; 1.  
 DR SEQUENCE 628 AA; 69195 MW; 9B01CB5ADECEAC77 CRC64;  
 Query Match 19.3%; Score 72.5; DB 1; Length 628;  
 Best Local Similarity 29.6%; Pred. No. 9.8;





```

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
reflected and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See http://www.isb-sib.ch/announcement/
or send an email to licens@isb-sib.ch).
-----
EMBL: U89277; AAC51169.1;
EMBL: BC002871; AAHG2871.1; ALT_INIT.
EMBL: BC017748; AAH17748.1; ALT_INIT.
Genew: BCNC3182; PBC1.
MIM: 602378;
GO: GO:0005634; Cytosol; TAS.
GO: GO:000515; Fibroblast binding activity; TAS.
InterPro: IPR001663; SAM.
Pfam: PF00546; SAM; 1.
SMART: SM00454; SAM; 1.
PROSITE: PSF105; SAM_ICMAIN; 1.
Developmental protein; DNA-binding; Nuclear protein; Metal-binding;
Zinc-finger.
425 441 POLY-GLN.
DOMAIN 940 1004 SAM.
ZNFINE 800 823 C4-TYPE (POTENTIAL).
CONFLICT 568 569 HL -> LK (IN REF. 1).
CONFLICT 613 613 S -> T (IN REF. 1).
CONFLICT 693 693 A -> T (IN REF. 1).
CONFLICT 752 752 V -> G (IN REF. 1).
CONFLICT 782 782 L -> F (IN REF. 1).
CONFLICT 972 972 L -> F (IN REF. 1).
SEQUENCE 1004 AA: 105418.NW: 5E3576575P904CAF CRC64;
Query Match 19.0%; Score 71.5; DB 1; Length 1004;
Fast Local Similarity 45.88; Pred. No. 21;
Matches 24; Conservative 5; Mismatches 23; Indels 15; Gaps 3;
OY 2 PPAPOPVDSIQVSSQSGQAVTVSRQPSGLINAYNSLTRSGLKRTTP-SLKPPVPPKPSFAP 60
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 448 PCPPQVPTQVPPSCSGQACTLVQPMQSG-----SPLSLPPDAAKRP---P 493
OY 51 LSTSMKP 67
DE 494 IPQSKP 500
Search completed: September 30, 2003, 16:33:00
Job time : 40 secs

```

GenCore version 5.1.6  
Copyright: (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: September 30, 2003, 16:19:53 : Search time 69 seconds  
(without alignments)  
269.272 Million cell updates/sec

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAPQVRVDSIQVHSSQPSG.....PFRKPSFALSTSMKPNDACT 72

Scoring table: BL2SUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 254552604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:

- 1: sp\_archaea:
- 2: sp\_bacteria:
- 3: sp\_fungi:
- 4: sp\_human:
- 5: sp\_invertebrate:
- 6: sp\_mammal:
- 7: sp\_mnc:
- 8: sp\_organelle:
- 9: sp\_phage:
- 10: sp\_plant:
- 11: sp\_rodent:
- 12: sp\_virus:
- 13: sp\_vertebrate:
- 14: sp\_unclassified:
- 15: sp\_virus:
- 16: sp\_bacteria:
- 17: sp\_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	376	100.0	507	Q96T04	Q96T04 homo sapien
2	376	100.0	562	Q96SY4	Q96SY4 homo sapien
3	376	100.0	562	Q8NC49	Q8NC49 homo sapien
4	376	100.0	574	Q96SM8	Q96SM8 homo sapien
5	376	100.0	699	Q96SM4	Q96SM4 homo sapien
6	376	100.0	1005	Q96Q71	Q96Q71 mus musculu
7	376	100.0	1030	Q96ZP5	Q96ZP5 homo sapien
8	376	100.0	1049	Q962H9	Q962H9 homo sapien
9	163.5	43.5	416	Q95KA6	Q95KA6 ma-aca fasc
10	163.5	43.5	451	Q96HK4	Q96HK4 homo sapien
11	163.5	43.5	454	Q96HG5	Q96HG5 homo sapien
12	163.5	43.5	998	Q8NFY6	Q8NFY6 homo sapien
13	163.5	43.5	1011	Q8NFY3	Q8NFY3 homo sapien
14	163.5	43.5	1017	Q8NFY5	Q8NFY5 homo sapien
15	163.5	43.5	1022	Q9P249	Q9P249 homo sapien
16	163.5	43.5	1073	Q8NFY4	Q8NFY4 homo sapien

#### ALIGNMENTS

RESULT :

Q96T04 PRELIMINARY: PRI: 507 AA.  
 ID Q96T04  
 A: Q96T04  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DZ 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein FLJ14533.  
 OS Homo sapiens (Human).  
 EC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 SC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Watanabe S., Hosoki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Katanabe S., Kimura K., Marakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Nimomiya K., Iwayanagi T.  
 RT \*NEDO human cDNA sequencing project.\*  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK027439; BAB55111.1;  
 DR InterPro; IPR003659; Plexin-like.  
 DR SMART; SM00423; PSI: 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 507 AA; 55464 MW; 80C567B438C51B39 CRC64;

Query Match 100.0%; Score 376; DS 4; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-33;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQVRVDSIQVHSSQPSGQAVTVSQPSLNATNSLTRSLKTPSLKPDVPPKPSFAP 60

2b 436 PPPAPQVRVDSIQVHSSQPSGQAVTVSQPSLNATNSLTRSLKTPSLKPDVPPKPSFAP 495

Qy 61 LSTSMKPNDACT 72

|||||

DB 496 LSTSMKPNDACT 562

RESULT 2  
Q96SY4

ID Q96SY4 PRELIMINARY: PRT: 562 AA.  
AC Q96SY4 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Hypothetical protein FLJ14565.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagaatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.,  
RI "NEDO human cDNA sequencing project."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK027471; BAB5136.1;  
DR InterPro: IPR003659; Plexin-like.  
DR SMART: SM00423; PS2.1;  
KW Hypothetical protein.  
SQ SEQUENCE 562 AA: 6133 MW: 6AB36957A7D578A CRC64:

Query Match 100.0% Score 376; DB 4; Length 562;  
Best Local Similarity 100.0%; Pred. No. 6.3e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPAPQVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 60  
DB 491 PPAPQVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 550  
QY 61 LSTSMKPNDACT 72  
DB 551 LSTSMKPNDACT 562

RESULT 3

ID Q96SY4 PRELIMINARY: PRT: 562 AA.  
AC Q96SY4 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DE 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
DE Hypothetical protein FLJ50494.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,  
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kodaira H., Sudo H.,  
RA Wagaatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Kojima S., Nagahari K.,  
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.,  
RI "NEDO human cDNA sequencing project."  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK074575; BAB1126.1;  
KW Hypothetical protein.  
SQ SEQUENCE 562 AA: 61286 MW: 709041459E34578A CRC64:

Query Match 100.0% Score 376; DB 4; Length 562;  
Best Local Similarity 100.0%; Pred. No. 6.3e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPAPQVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 60  
DB 491 PPAPQVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 550  
QY 61 LSTSMKPNDACT 72  
DB 551 LSTSMKPNDACT 562

RESULT 4

ID Q96SY4 PRELIMINARY: PRT: 574 AA.  
AC Q96SY4 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Hypothetical protein FLJ14746.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagaatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.,  
RI "NEDO human cDNA sequencing project."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK027654; BAB5265.1;  
DR InterPro: IPR003659; Plexin-like.  
DR SMART: SM00423; PS2.1;  
KW Hypothetical protein.  
SQ SEQUENCE 574 AA: 62822 MW: 0C79E01A4117A95 CRC64:

Query Match 100.0% Score 376; DB 4; Length 574;  
Best Local Similarity 100.0%; Pred. No. 6.5e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPAPQVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 60  
DB 503 PPAPQVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 562  
QY 61 LSTSMKPNDACT 72  
DB 563 LSTSMKPNDACT 574

RESULT 5

ID Q96SY4 PRELIMINARY: PRT: 599 AA.  
AC Q96SY4 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Hypothetical protein FLJ14565.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagaatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.,

```
RT "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK027501; BAB55158.1;
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; SEMA.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
KW Hypothetical protein.
SQ SEQUENCE 599 AA; 76723 MW; 2E5F11D59741394 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 628 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 607
QY 61 LSTSMKPNDACT 72
DB 688 LSTSMKPNDACT 699

Query Match 100.0%; Score 376; DB 4; Length 1030;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 959 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1618
QY 61 LSTSMKPNDACT 72
DB 1019 LSTSMKPNDACT 1030

RESULT 8
Q9H2H9 PRELIMINARY; PRT; 1049 AA.
ID Q9H2H9
AC Q9H2H9
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical protein KIAA1368 (Fragment).
GN KIAA1368
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20564339; PubMed=10993894;
PA Nagase T., Kikuno R., Ishikawa K., Hirose M., Chazawa M., Chazawa M.,
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL: A037789; BAA92606.1;
DR Genes; HGNC:10738; SEMA6A;
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001527; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00530; Sema; 1.
KW Hypothetical protein.
FT NON-TER
SQ SEQUENCE 1045 AA; 114511 MW; 7781D20ACC7A9AA CRC64;

Query Match 100.0%; Score 376; DB 4; Length 1049;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 959 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1618
QY 61 LSTSMKPNDACT 72
DB 1019 LSTSMKPNDACT 1030

RESULT 7
Q9H2E6 PRELIMINARY; PRT; 1030 AA.
ID Q9H2E6
AC Q9H2E6
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Hypothetical protein.
GN KIAA1368
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20564339; PubMed=10993894;
PA Nagase T., Kikuno R., Ishikawa K., Hirose M., Chazawa M., Chazawa M.,
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL: A037789; BAA92606.1;
DR Genes; HGNC:10738; SEMA6A;
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001527; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00530; Sema; 1.
KW Hypothetical protein.
FT NON-TER
SQ SEQUENCE 1005 AA; 579699277458079D CRC64;

Query Match 100.0%; Score 376; DB 11; Length 1005;
Best Local Similarity 100.0%; Pred. No. 1.2e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 914 PPPAPQVDSIOVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 993
QY 61 LSTSMKPNDACT 72
DB 994 LSTSMKPNDACT 1005

RESULT 7
Q9H2E6 PRELIMINARY; PRT; 1030 AA.
ID Q9H2E6
AC Q9H2E6
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Hypothetical protein.
GN KIAA1368
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20564339; PubMed=10993894;
PA Nagase T., Kikuno R., Ishikawa K., Hirose M., Chazawa M., Chazawa M.,
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL: A037789; BAA92606.1;
DR Genes; HGNC:10738; SEMA6A;
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001527; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00530; Sema; 1.
KW Hypothetical protein.
FT NON-TER
SQ SEQUENCE 1005 AA; 579699277458079D CRC64;
```



```

DR SMART: SMO0630; Sema: 1.
SQ SEQUENCE 998 AA: 111760 MW: 3F46D6872EB9D5544 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 998;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRSIQVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 927 PPTGAKVDYIQ-----GTPVSVHLQPSLSKSSSYTSNGTLPRTGKRTPSLKPDVPP 979
QY 55 KPSFAPLSTSMKP 67
DB 930 KPSEVPQTPSVRP 992

RESULT 13
Q8NFY3 PRELIMINARY: PRT: 1017 AA.
AC Q8NFY3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22; Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23; Last annotation update)
DE Semaphorin 6D isoform 1.
GN SEMA6D.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ou X., Wei H., Zhai Y., Yu Y., Tang F., He F.;
RT "Homo sapiens semaphorin 6D isoform 1 (SEMA6D); cDNA, complete cds."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF389430; AA069453.1;
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SMO0630; Sema: 1.
SQ SEQUENCE 1011 AA: 13269 MW: 9D6B8B3631941589 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 1011;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRSIQVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 940 PPTGAKVDYIQ-----GTPVSVHLQPSLSKSSSYTSNGTLPRTGKRTPSLKPDVPP 992
QY 55 KPSFAPLSTSMKP 67
DB 993 KPSEVPQTPSVRP 1005

RESULT 14
Q8NFY5 PRELIMINARY: PRT: 1017 AA.
AC Q8NFY5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22; Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23; Last annotation update)
DE Semaphorin 6D isoform 3.
GN SEMA6D.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ou X., Wei H., Zhai Y., Yu Y., Tang F., He F.;
RT "Homo sapiens semaphorin 6D isoform 3 (SEMA6D); cDNA, complete cds."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL: AF389428; AA069451.1;
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SMO0630; Sema: 1.
SQ SEQUENCE 1017 AA: 133736 MW: 4D639CEBADD9F2A0 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 1017;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRSIQVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 946 PPTGAKVDYIQ-----GTPVSVHLQPSLSKSSSYTSNGTLPRTGKRTPSLKPDVPP 998
QY 55 KPSFAPLSTSMKP 67
DB 999 KPSEVPQTPSVRP 1011

RESULT 15
Q8Q249 PRELIMINARY: PRT: 1022 AA.
AC Q8Q249
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15; Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 24; Last annotation update)
DE Hypothetical protein K1AA1479 (Fragment).
GN K1AA1479.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20277182; PubMed=8819331;
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human
RT genes XVII. The complete sequences of 100 new cDNA clones from brain
RT which code for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
DR EMBL: AR040912; BA055003.2;
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SMO0424; Sema; 1.
DR SMART: SMO0630; Sema; 1.
KW Hypothetical protein.
FT NON-TER 1
SQ SEQUENCE 1022 AA: 114372 MW: BE4FRD5EA02C59C4 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 1022;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRSIQVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 951 PPTGAKVDYIQ-----GTPVSVHLQPSLSKSSSYTSNGTLPRTGKRTPSLKPDVPP 1003
QY 55 KPSFAPLSTSMKP 67
DB 1004 KPSEVPQTPSVRP 1016

Search completed: September 30, 2003, 16:34:26
Job time : 75 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer, Ltd.

CM protein - protein search, using sw model

Run on: September 30, 2003, 16:31:08 : Search time 42 seconds  
(without alignments)  
72.533 Million cell updates/sec

Title: us-09-856-681-4

Perfect score: 376

Sequence: 1 PEPAPQVDSIQVHSSQPSG.....PPKPSAPLSTSMKPNDACT 72

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 128717 seqs, 4210858 residues

Total number of hits satisfying chosen parameters: 128717

Minimum DB seq length: 0

Maximum DB seq length: 226000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued\_Patents\_AA:\*

1: /cqn2\_5/ptdata/1/iaa/5A\_COMB.pep.\*

2: /cqn2\_5/ptdata/1/iaa/5B\_COMB.pep.\*

3: /cqn2\_5/ptdata/1/iaa/5A\_COMB.pep.\*

4: /cqn2\_5/ptdata/1/iaa/5B\_COMB.pep.\*

5: /cqn2\_5/ptdata/1/iaa/5A\_COMB.pep.\*

6: /cqn2\_5/ptdata/1/iaa/5B\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	72	19.1	480	3	US-09-189-035-5
2	72	19.1	480	3	US-09-382-086-5
3	71.5	19.0	1004	3	US-08-916-352-2
4	69.5	18.5	1125	4	US-08-913-783A-152
5	69.5	18.5	1610	4	US-09-513-783A-22
6	69	18.4	400	4	US-09-252-991A-32313
7	68.5	18.2	525	4	US-09-196-270-6
8	68	18.1	1149	3	US-08-560-005-5
9	68	18.1	1149	3	US-08-416-540-5
10	68	18.1	1149	4	US-09-969-528-5
11	68	18.1	1253	1	US-08-252-956B-12
12	68	18.1	1261	1	US-08-252-956B-18
13	67	17.8	169	4	US-09-252-991A-32563
14	67	17.8	2035	1	US-08-046-585-5
15	67	17.8	2035	1	US-08-393-703-5
16	67	17.8	2035	5	PCT-US93-11721-5
17	65.5	17.4	167	4	US-09-252-991A-32720
18	65.5	17.4	366	4	US-09-252-991A-32385
19	65.5	17.4	1142	2	US-08-993-116-7
20	65.5	17.4	1142	3	US-08-845-528C-7
21	65.5	17.4	1142	3	US-09-061-709-2
22	65.5	17.4	1142	4	US-09-566-2815-7
23	65.5	17.4	1422	4	US-09-499-651-2
24	65.5	17.4	1297	3	US-09-540-245A-17
25	65	17.3	70	4	US-09-006-428A-15
26	65	17.3	143	4	US-09-252-991A-22566
27	65	17.3	351	3	US-08-466-465-6

28	64.5	17.2	433	3	US-09-046-158A-2
29	64.5	17.2	897	1	US-07-960-389-2
30	64.5	17.2	1069	4	US-09-233-857-13
31	64.5	17.2	1099	4	US-09-442-100-2
32	64.5	17.2	1099	4	US-08-939-106-2
33	64.5	17.2	3969	3	US-08-061-376-5
34	64	17.0	641	4	US-09-071-035-456
35	64	17.0	1313	4	US-09-071-035-450
36	64	17.0	1313	4	US-09-071-035-454
37	63.5	16.9	280	4	US-09-252-991A-20783
38	63.5	16.9	302	4	US-09-252-991A-21231
39	63.5	16.9	479	4	US-09-252-991A-32594
40	63	16.8	816	2	US-08-785-310A-8
41	63	16.8	816	2	US-08-816-893A-53
42	63	16.8	816	3	US-08-885-291-53
43	63	16.8	816	3	US-09-496-672-53
44	62.5	16.6	846	4	US-09-858-664A-3
45	62	16.5	304	4	US-09-493-343B-22

#### ALIGNMENTS

#### RESULT 1

US-09-189-035-5

: Sequence 5, Application US/09189035

: Patent No. 6020165

: GENERAL INFORMATION:

: APPLICANT: Yue, Henry

: APPLICANT: Corley, Neil C.

: APPLICANT: Guegler, Karl J.

: APPLICANT: Baughn, Mariah R.

: TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS

: FILE REFERENCE: PP-0638 US

: CURRENT APPLICATION NUMBER: US/09/189,035

: CURRENT FILING DATE: 1998 11-10

: NUMBER OF SEQ ID NOS: 6

: SOFTWARE: PERL Program

: SEQ ID NO 5

: LENGTH: 460

: TYPE: PRT

: ORGANISM: Homo sapiens

: FEATURE: -

: OTHER INFORMATION: 32245671

US-09-189-035-5

Query Match	19.1%	Score 72	DB 3	Length 480
Best Local Similarity	31.2%	Pred. No. 4.8		
Matches	24	Conservative	13	Mismatches
			26	Indels
			14	Gaps
			4	
QY	1	PPAPQVDSIQVHSSQPSGQAVTSRPSLNAYNSLTRS-----GLKRTPSL-----KP 50		
DB	264	PTPSPPIIG--PAPCSAPSSQYGTMTIQ--ISRHNSTTSSTSSGGYRRTFTSVTAQPSAQ 319		
QY	51	DVPPKPSAPLSTSMKP 67		
DB	320	IVNGSPYQSNSISAP 336		

#### RESULT 2

US-09-382-086-5

: Sequence 5, Application US/09382086

: Patent No. 620106

: GENERAL INFORMATION:

: APPLICANT: Yue, Henry

: APPLICANT: Corley, Neil C.

: APPLICANT: Guegler, Karl J.

: APPLICANT: Baughn, Mariah R.

: TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS

: FILE REFERENCE: PP-0638 US

: CURRENT APPLICATION NUMBER: US/09/382,086

: CURRENT FILING DATE: 1999-08-24

: EARLIER APPLICATION NUMBER: 09/189,035

1 EARLIER FILING DATE: 1998-11-10  
2 NUMBER OF SEQ ID NOS: 6  
3 SOFTWARE: PERL Program  
4 SEQ ID NO 5  
5 LENGTH: 480  
6 TYPE: PRT  
7 ORGANISM: Homo sapiens  
8 FEATURE:   
9 OTHER INFORMATION: g245671  
10 US-09-382-086-5

Query Match 19.1%; Score 71.5; DB 3; Length 1004;  
Best Local Similarity 31.2%; Pred. No. 4; 3;  
Matches 24; Conservative 13; Mismatches 26; Indels 14; Gaps 4;  
QY 1 PPAQVDSIQVHSSQSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 60  
DB 264 PTPSPPTIG--PAGSAGSGQVGTMC--ISRHNSTTS--SSGGRYRTSVTAQPSAQP 429  
QY 51 DVPKPSFAPLSTSKP 67  
DB 320 RVNGGPLYQNSISLAP 136

RESULT 3  
US-08-916-352-2  
1 Sequence 2, Application US/08916352  
2 Patent No. 6166191  
3 GENERAL INFORMATION:  
4 APPLICANT: CHIRON CORPORATION  
5 TITLE OF INVENTION: HUMAN FOLYHOMEOGIC 1 (hpl1) ACTS AS A  
6 NUMBER OF SEQUENCES: 2  
7 CORRESPONDENCE ADDRESS:  
8 ADDRESSEE: CHIRON CORPORATION  
9 STREET: 4560 HORTON STREET  
10 CITY: EMERYVILLE  
11 STATE: CA  
12 COUNTRY: USA  
13 ZIP: 94608  
14 COMPUTER READABLE FORM:  
15 MEDIUM TYPE: floppy disk  
16 COMPUTER: IBM PC compatible  
17 OPERATING SYSTEM: PC-DOS/MS-DOS  
18 SOFTWARE: Patent Release #1.0, Version #1.0  
19 CURRENT APPLICATION DATA:  
20 APPLICATION NUMBER: US/06/316,352  
21 FILING DATE:  
22 CLASSIFICATION: 435  
23 ATTORNEY/AGENT INFORMATION:  
24 NAME: FOTTER, JANE  
25 REGISTRATION NUMBER: 33,332  
26 REFERENCE/DOCKET NUMBER: 1355.  
27 TELECOMMUNICATION INFORMATION:  
28 TELEPHONE: 510-921-2707  
29 TELEFAX: 510-655-3542  
30 INFORMATION FOR SEQ ID NO: 2:  
31 SEQUENCE CHARACTERISTICS:  
32 LENGTH: 1004 amino acids  
33 TYPE: amino acid  
34 STRANDEDNESS: single  
35 TOPOLOGY: linear  
36 MOLECULE TYPE: protein  
37 US-08-916-352-2

Query Match 19.0%; Score 71.5; DB 3; Length 1004;  
Best Local Similarity 35.8%; Pred. No. 14;  
Matches 24; Conservative 5; Mismatches 23; Indels 15; Gaps 3;  
QY 2 PPAQVDSIQVHSSQSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 60  
DB 448 POPQVPTQQVPPSGSQDAGTAVQPMLOS-----SPLSLPPDAAKPK---P 493

QY 51 LSTSKAP 67  
DB 494 PPIQSKP 500  
RESULT 4  
US-09-513-783A-152  
1 Sequence 22, Application US/09513783A  
2 Patent No. 6416959  
3 GENERAL INFORMATION:  
4 APPLICANT: Giuliano, Kenneth A.  
5 TITLE OF INVENTION: A System for Cell Based Screening  
6 FILE REFERENCE: 97-022-11  
7 CURRENT APPLICATION NUMBER: US/09/513,783A  
8 CURRENT FILING DATE: 2000-02-25  
9 NUMBER OF SEQ ID NOS: 180  
10 SOFTWARE: PatentIn Ver. 2.0  
11 SEQ ID NO 152  
12 LENGTH: 1125  
13 TYPE: PRT  
14 ORGANISM: Mus musculus  
15 US-09-513-783A-152

Query Match 18.5%; Score 69.5; DB 4; Length 1125;  
Best Local Similarity 28.6%; Pred. No. 26;  
Matches 18; Conservative 13; Mismatches 31; Indels 1; Gaps 1;  
QY 3 PAP-QRVDISQVHSSQSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61  
DB 516 PPSLENLEOKETPGSQPSQPSQVSRQAEAKAAGVTGNDITTPNKEPPSPKAKPL 675  
QY 52 STS 64  
DB 676 ATT 678

RESULT 5  
US-09-513-783A-22  
1 Sequence 22, Application US/09513783A  
2 Patent No. 6416959  
3 GENERAL INFORMATION:  
4 APPLICANT: Giuliano, Kenneth A.  
5 TITLE OF INVENTION: A System for Cell Based Screening  
6 FILE REFERENCE: 97-022-11  
7 CURRENT APPLICATION NUMBER: US/09/513,783A  
8 CURRENT FILING DATE: 2000-02-25  
9 NUMBER OF SEQ ID NOS: 180  
10 SOFTWARE: PatentIn Ver. 2.0  
11 SEQ ID NO 22  
12 LENGTH: 1610  
13 TYPE: PRT  
14 ORGANISM: Artificial Sequence  
15 FEATURES:  
16 OTHER INFORMATION: Description of Artificial Sequence:  
17 OTHER INFORMATION: EYFF-DEVD-MAP4-EBFP construct  
18 US-09-513-783A-22

Query Match 18.5%; Score 69.5; DB 4; Length 1610;  
Best Local Similarity 28.6%; Pred. No. 41;  
Matches 18; Conservative 13; Mismatches 31; Indels 1; Gaps 1;  
QY 3 PAP-QRVDISQVHSSQSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61  
DB 552 PPSLENLEOKETPGSQPSQVSRQAEAKAAGVTGNDITTPNKEPPSPKAKPL 921  
QY 62 STS 64  
DB 922 ATT 924

RESULT 6



```
US-09-252-931A-32313
: Sequence 32313, Application US/09252991A
: Patent No. 6551755
: GENERAL INFORMATION:
: APPLICANT: Marc J. R-benfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.135
: CURRENT APPLICATION NUMBER: US/09/252.991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: 65 60/274.788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: 65 60/294.190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 32313
: LENGTH: 400
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32313

Query Match: 18.4%; Score 69; DB 4; Length 400;
Best Local Similarity: 30.1%; Pred. No. 812;
Matches 25; Conservative 9; Mismatches 20; Indels 24; Gaps 5;

QY 2 PPAPQV-----DSIQVHSSQPSGAVTVSRQPSLNAYNSLTRSGIKRTP 47
DB 115 PPATPRRLRRGGRAGATGCGASAGCAP-GCAVINGAPA---GAEFTAG--RAH 145

QY 48 LKPPVPKPSFAPLSTSMKINDA 70
DB 166 NRPOPPATSESPMS---KTPA 185

RESULT 7
US-09-196-270-6
: Sequence 6, Application US/09196270
: Patent No. 650636
: GENERAL INFORMATION:
: APPLICANT: Hecht, Peter
: APPLICANT: Madden, Kevin
: APPLICANT: Fink, Gerald
: TITLE OF INVENTION: CHIMERIC PRE-ACTIVATED TRANSCRIPTION
: TITLE OF INVENTION: FACTORS
: FILE REFERENCE: 50078/004002
: CURRENT APPLICATION NUMBER: US/09/196,270
: CURRENT FILING DATE: 1998-11-19
: EARLIER APPLICATION NUMBER: 60/066,129
: EARLIER FILING DATE: 1997-11-19
: EARLIER APPLICATION NUMBER: 60/066,308
: EARLIER FILING DATE: 1997-11-21
: EARLIER APPLICATION NUMBER: 60/066,462
: EARLIER FILING DATE: 1997-11-24
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 625
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
US-09-196-270-6

Query Match: 16.2%; Score 68.5; DB 4; Length 625;
Best Local Similarity: 25.8%; Pred. No. 16;
Matches 31; Conservative 13; Mismatches 49; Indels 49; Gaps 6;

QY 1 PPAPQVQVDSIQVHSS-----QPSQGVV-----VSRQ-24
DB 418 PPNAPS-YOVQCGSSSATANTATYVVRVLRATPTGPTTEMLPNHSHNAGGVNROS 476

QY 29 -----PCLNAYNSLTRSGIKRTPSLKPPVPPKSPAPES--TSMKPN---GAVT 72
DB 477 QVAMPHYISVRAATSYSSSSGSLPPVQSKIPN:FSRRINAKOTS:KPNWEEISLNKSKSI 546
```

```
RESULT 8
US-08-560-005-5
: Sequence 5, Application US/08560005
: Patent No. 6001354
: GENERAL INFORMATION:
: APPLICANT: Pot, David A.
: APPLICANT: Williams, Lewis T.
: APPLICANT: Jefferson, Anne Bennett
: APPLICANT: Majerus, Philip W.
: TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
: TITLE OF INVENTION: Acids Encoding Therefor
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew
: STREET: One Market Plaza, Steuart Tower, Suite 2000
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/560,005
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Dow, Karen H.
: REGISTRATION NUMBER: 25,664
: REFERENCE/DOCKET NUMBER: 2307K-0624000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-326-2400
: TELEFAX: 415-326-2422
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 144 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Region
: LOCATION: 1..145
: OTHER INFORMATION: /note="51c"
US-08-560-005-5

Query Match: 18.1%; Score 68; DB 3; Length 1149;
Best Local Similarity: 29.9%; Pred. No. 40;
Matches 23; Conservative 12; Mismatches 16; Indels 5; Gaps 3;

QY 2 PPAPQV---DSIQVHSSQPSG---QAVTVSRQPSLNAYNSLTRSGIKRTP-SLKPPVPPK 55
DB 380 PPAPRAAPRSEPTPRKPEGAPEGVAAAPPKPSFNFPAYVYLVGVPHQLPPPPPS 939

QY 56 PSFAPLSTSMKPNDACT 72
DB 540 PARAPVPSATKKNVAIT 956

RESULT 9
US-09-418-940-5
: Sequence 5, Application US/09418540
: Patent No. 6296848
: GENERAL INFORMATION:
: APPLICANT: Pot, David A.
: APPLICANT: Williams, Lewis T.
: APPLICANT: Jefferson, Anne Bennett
: APPLICANT: Majerus, Philip W.
: TITLE OF INVENTION: No. 6296848el Grb2 Associating Protein and Nucleic
: TITLE OF INVENTION: Acids Encoding Therefor
: NUMBER OF SEQUENCES: 10
```

```

CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/418-540
FILING DATE: 14-OCT-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/560,005
FILING DATE: 17-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0624000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2422
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Region
LOCATION: 1..1149
OTHER INFORMATION: /note= "51c"
US-09-418-540-5

Query Match 18.1%; Score 68; DB 3; Length 1149;
Best Local Similarity 29.9%; Pred. No. 40;
Matches 23; Conservative 12; Mismatches 36; Indels 6; Gaps 3;

QY 2 PPAQGV--DSIQVHSSCPG--QAVTVSRPSLNAYNSITRSGLEKHTP-SLKPQVPPK 55
DQ 880 PPAAPRAAPREPLTPRLKPEGAPEPGVAAPPKNSFNPNPAYVYEGVPHLLPPEPPS 939
QY 56 PSFAPLSTSMKPNCACT 72
DQ 940 PARAPVPSATKNKVAIT 956

RESULT 10
US-09-969-528-5
Sequence 5, Application US/09/969528
Patent No. 5472197
GENERAL INFORMATION:
APPLICANT: Pot, David A.
Williams, Lewis T.
Jefferson, Anne Bennett
Majerus, Philip W.
TITLE OF INVENTION: No. 6472197el 31b2 Association Protein and Nucleic
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/969,528
FILING DATE: 01-OCT-2001
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/560,005
FILING DATE: <UNKNOWN>
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0624000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2430
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Region
LOCATION: 1..1149
OTHER INFORMATION: /note= "51c"
US-09-969-528-5

Query Match 18.1%; Score 68; DB 4; Length 1149;
Best Local Similarity 29.9%; Pred. No. 40;
Matches 23; Conservative 12; Mismatches 36; Indels 6; Gaps 3;

QY 2 PPAQGV--DSIQVHSSCPG--QAVTVSRPSLNAYNSITRSGLEKHTP-SLKPQVPPK 55
DQ 880 PPAAPRAAPREPLTPRLKPEGAPEPGVAAPPKNSFNPNPAYVYEGVPHLLPPEPPS 939
QY 56 PSFAPLSTSMKPNCACT 72
DQ 940 PARAPVPSATKNKVAIT 956

RESULT 11
US-08-252-966B-12
Sequence 12, Application US/08252966B
Patent No. 5624818
GENERAL INFORMATION:
APPLICANT: Eisenman, Robert N.
APPLICANT: Hufn, Peter J.
APPLICANT: Aver, Donald E.
TITLE OF INVENTION: Regulatory Proteins that Dimerize with
TITLE OF INVENTION: Mad or Max
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSES: Christensen, O'Connor, Johnson, and Kindness PLLC
STREET: 1420 Fifth Ave., Suite 2800
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,966B
FILING DATE: 01-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: FHCRI7694

```

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 682-8100  
TELEFAX: (206) 224-0779

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1253 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

DESCRIPTION: translation of msina cDNA; see Figure 23

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Mus musculus

US-08-252-9668-12

Query Match 18.1% Score 68; DB 1; Length 1253;  
Best Local Similarity 30.7%; Pred. No. 44;  
Matches 23; Conservative 7; Mismatches 35; Indels 10; Gaps 3;

QY 1 PPAPOVDISIQVHSSQPSQAVTVSRQPS--LNAYNSLIRSLKRTISLKDVPKPSF 58  
DB 238 PQPPFQ-----HPSQSSQSNAPTPAQAPQPTAAKYKPSQLOAHTPASQQTPLPPY 290  
QY 59 A-PLSTSMKPNDACT 72  
DB 291 ASRSPVPVQPHPTVT 305

## RESULT 12

US-08-252-9668-14

Sequence 16, Application US/082529668

Patent No. 5624818

GENERAL INFORMATION:

APPLICANT: Eisenman, Robert N.

APPLICANT: Hurlin, Peter J.

APPLICANT: Ayer, Donald F.

TITLE OF INVENTION: Regulatory Proteins that Dimorize with

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Christensen, Gordon, Johnson, and Kindness PLLC

STREET: 1420 Fifth Ave., Suite 2800

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98101-2347

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/03/252,966B

FILING DATE: 01-JUN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Shelton, Dennis K.

REGISTRATION NUMBER: 26,357.

REFERENCE/DOCKET NUMBER: FHC817694

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 682-8100

TELEFAX: (206) 224-0779

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 1261 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

DESCRIPTION: translation of msina9 cDNA; see Figure 29A, B, C, D

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Mus musculus

US-08-252-9668-18

Query Match 18.1% Score 68; DB 1; Length 1261;

Best Local Similarity 30.7%; Pred. No. 44;

Matches 23; Conservative 7; Mismatches 35; Indels 10; Gaps 3;

QY 1 PPAPOVDISIQVHSSQPSQAVTVSRQPS--LNAYNSLIRSLKRTISLKDVPKPSF 58  
DB 238 PQPPFQ-----HPSQSSQSNAPTPAQAPQPTAAKYKPSQLOAHTPASQQTPLPPY 290  
QY 59 A-PLSTSMKPNDACT 72  
DB 291 ASRSPVPVQPHPTVT 305

## RESULT 13

US-08-252-991A-30563

Sequence 30563, Application US/09252991A

Patent No. 6351795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubinfeld et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107136.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-15

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 30563

LENGTH: 169

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30563

Query Match 17.8% Score 67; DB 4; Length 169;

Best Local Similarity 27.4%; Pred. No. 4.6;

Matches 23; Conservative 14; Mismatches 29; Indels 18; Gaps 3;

QY 3 PAPAQVDSIQVHSSQPSQAVTVSRQPSLNAYNSL-----IRSLKRTPSLK---- 49  
DB 2 PQSRATSTPSSSSSPARIACSGCKKNPNRAGSTLPAPSPCCSRSRQSRSTRTCG 61  
QY 50 --PVPKPSAPLSTSMKPNDACT 71  
DB 62 WTPTVPVPA---ASTSIRPTRKC 82

## RESULT 14

US-08-046-585-5

Sequence 5, Application US/08046585

Patent No. 5453362

GENERAL INFORMATION:

APPLICANT: Lemarco, Kelly

APPLICANT: Wilson, Angus

APPLICANT: Heitz, Winship

TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:

TITLE OF INVENTION: HOST CELL FACTOR

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, ROHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/24/046,585

FILING DATE: 12-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57503-1/RAG

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 5

SEQUENCE CHARACTERISTICS:

LENGTH: 2035 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-046-585-5

Query Match 17.8%; Score 67; DB 1; Length 2035;

Best Local Similarity 32.8%; Pred. No. 1e+02;

Matches 22; Conservative 10; Mismatches 27; Indels 8; Gaps 3;

QY 2 PPAPQVDSIQVHSQPSG-----QAVTVSRQPSLNAYNSLTR--SGLKRTPSLKPDVPP 54

Db 1741 PSTVALLPSTATSLAPSNITVAPOPVVVASPAKIQAAALTEVANGIE-SLGVKPDIPP 1799

QY 55 KPSFAPL 61

Db 1800 PPSKAPM 1806

## RESULT 15

US-08-393-703-5

Sequence 5, Application US/08393703

Patent No. 5585239

GENERAL INFORMATION:

APPLICANT: Lemarco, Kelly

APPLICANT: Wilson, Angus

APPLICANT: Hertz, Wilsch

TITLE OF INVENTION: A NOVEL PHARYNGEAL TRANSCRIPTION PROTEIN;

TITLE OF INVENTION: HOST CELL FACTOR

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEER, HORBACH, TEST, ALBERTSON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/28/393,703

FILING DATE: 24-FEB-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57503-2/RAG

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2035 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-393-703-5

Query Match 17.8%; Score 67; DB 1; Length 2035;

Best Local Similarity 32.8%; Pred. No. 1e+02;

Matches 22; Conservative 10; Mismatches 27; Indels 8; Gaps 3;

QY 2 PPAPQVDSIQVHSQPSG-----QAVTVSRQPSLNAYNSLTR--SGLKRTPSLKPDVPP 54

Db 1741 PSTVALLPSTATSLAPSNITVAPOPVVVASPAKIQAAALTEVANGIE-SLGVKPDIPP 1799

QY 55 KPSFAPL 61

Db 1800 PPSKAPM 1806

Search completed: September 30, 2003, 16:36:16

Job time : 43 secs